

79280

STIC-Biotech/ChemLib

184707

Mg

From: Swope, Sheridan
Sent: Friday, April 07, 2006 7:23 PM
To: STIC-Biotech/ChemLib
Subject: 10/825,632

For 10/825,632 , pls search and interference search:

SID 1 against the NT and AA databases.

SID 2 against the NT and AA databases.

Thanks!

Sheridan Swope, Ph.D.
Primary Patent Examiner
AU 1656/Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Not 4/27/06
6/20/06

RECEIVED
APR 10 2006
STIC

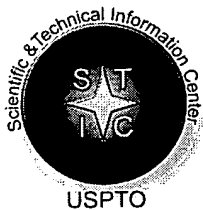
BEST AVAILABLE COPY

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 184707

TO: Sheridan Swope
Location: rem/2B71/3C70
Art Unit: 1656
Thursday, April 20, 2006
Case Serial Number: 10/825632

From: Kristine Hensle
Location: Biotech-Chem Library
REM-1B69
Phone: (571) 272-4161

Kristine. Hensle@uspto.gov

Search Notes

Examiner Swope,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Kristine Hensle
Librarian
STIC Biotech/Chem Library
(571)272-4161



THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 12:42:35 ; Search time 193 Seconds
(without alignments)
2007.938 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAAMETEGLGVIEFTADC.....HLLHYLQENLGSRIALKVI 882

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq 21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4700	100.0	882	4 AAB47187	AAB47187 Human DPP
2	4700	100.0	882	5 ABG61591	ABG61591 Human DPP
3	4700	100.0	882	5 AAG78415	AAG78415 Amino aci
4	4700	100.0	882	5 AAE24170	AAE24170 Human dip
5	4700	100.0	882	5 AAU74749	AAU74749 Human pro
6	4700	100.0	882	5 ADI17086	ADI17086 Human NOV
7	4700	100.0	882	5 ABU07720	ABU07720 Human ser
8	4528.5	96.4	883	5 ADI17085	ADI17085 Murine NO
9	4397.5	93.6	831	6 ABU92026	ABU92026 Human pro
10	4353	92.6	824	6 ABU92030	ABU92030 Human pro
11	4118	87.6	782	5 ABG97361	ABG97361 Novel hum
12	3901	83.0	746	8 ADI16334	ADI16334 Human pro
13	3817	81.2	738	8 ADT04044	ADT04044 Human pro
14	3771	80.2	724	5 ABG97362	ABG97362 Novel hum
15	3504	74.6	658	5 ABG61600	ABG61600 Human DPP
16	3504	74.6	661	5 ABG61596	ABG61596 Human DPP
17	3504	74.6	690	5 ABG61594	ABG61594 Human DPP
18	3333.5	70.9	632	4 AAB93565	AAB93565 Human pro
19	3236	68.9	613	5 ABG61601	ABG61601 Human DPP
20	3087	65.7	587	5 ADRA1398	ADRA1398 Human CD-
21	2870	61.1	863	5 ABG61592	ABG61592 Human DPP
22	2870	61.1	863	5 ADI17083	ADI17083 Human NOV
23	2870	61.1	892	5 ABG61602	ABG61602 Human DPP
24	2870	61.1	892	5 ABG61604	ABG61604 Human DPP

25	2870	61.1	892	5 ABB98134	ABB98134 Human PWM
26	2870	61.1	892	8 ADS10951	ADS10951 Human the
27	2863	60.9	969	5 AAE24168	AAE24168 Human dip
28	2862	60.9	863	5 ADI16690	ADI16690 Human NOV
29	2862	60.9	863	5 ADI16688	ADI16688 Human NOV
30	2862	60.9	863	8 ADN42344	ADN42344 Human nov
31	2862	60.9	863	8 ADN42342	ADN42342 Human nov
32	2835	60.3	830	5 AAE24171	AAE24171 Human dip
33	2833	60.3	869	5 AAE24169	AAE24169 Alternati
34	2820.5	60.0	879	5 ABG61607	ABG61607 Human DPP
35	2820.5	60.0	879	5 ABG61608	ABG61608 Human DPP
36	2820.5	60.0	879	8 ADS10953	ADS10953 Human the
37	2713.5	57.7	847	5 AAE23875	AAE23875 Murine di
38	2707	57.6	854	5 ADS10952	ADS10952 Human the
39	2671.5	56.8	873	8 ADS10479	ADS10479 Human the
40	2547.5	54.2	580	5 AAE14337	AAE14337 Human pro
41	2498	53.1	755	8 ADQ67811	ADQ67811 Novel hum
42	2422	51.5	465	4 AAB47189	AAB47189 Human DPP
43	2413.5	51.4	737	4 AAM38724	AAM38724 Human pol
44	2406	51.2	832	5 ABG61605	ABG61605 Human DPP
45	2406	51.2	832	5 ABG61606	ABG61606 Human DPP

ALIGNMENTS

RESULT 1

AAB47187

ID AAB47187 standard; protein; 882 AA.

XX

AC AAB47187;

XX

DT 29-JUN-2001 (first entry)

XX

DE Human DPP8.

XX

KW Human; dipeptidyl aminopeptidase; DPP8; prolol oligopeptidase;

KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;

KW growth hormone deficiency; glucose level; mucosal regeneration;

KW non-insulin dependent diabetes mellitus; glucose intolerance;

KW immunosuppression.

XX Homo sapiens.

OS

XX Key Location/Qualifiers

PH Active-site 739

FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

FT Active-site 817

FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

FT Active-site 849

FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX WO200119866-A1.

XX

PN 22-MAR-2001.

XX

PD 11-SEP-2000; 2000WO-AU001085.

XX

PR 10-SEP-1999; 99AU-00002762.

XX

PR 18-FEB-2000; 2000AU-00005709.

XX

PA (UNSY) UNIV SYDNEY.

XX

PI Abbott CA, Gorell MD;

XX

DR WPI; 2001-281520/29.

XX

DR N-PSDB; AAC85694.

XX

PT New human dipeptidyl aminopeptidase (DPP8) useful for cleaving

PT substrates, identifying inhibitors of DPP8 catalytic activity which have

PT therapeutic uses, and for detecting activated T cells.

XX

PS Claim 1; Fig 2; 78pp; English.

XX This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has
CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.
CC Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase,
CC because it is capable of hydrolyzing the peptide bond C-terminal to Pro
CC in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is
CC useful for cleaving a substrate, and for detecting an activated T cell
CC which involves measuring the level of DPP8 gene expression in a T cell.
CC The level of DPP8 expression is detected by detecting the amount of DPP8
CC RNA in the cell. It is also useful for identifying a molecule capable of
CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as
CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
CC growth hormone deficiency, lowering glucose levels in non-insulin
CC dependent diabetes mellitus and other disorders involving glucose
CC intolerance, enhancing mucosal regeneration and as immunosuppressants
XX
XX Sequence 882 AA;
SQ
Query Match 100.0%; Score 4700; DB 4; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMETQLGVPEITADCEENIESQDRPKLEPFYVERYSWSQLKLLADTRKHGYMM 60
DB 1 MAAAMETQLGVPEITADCEENIESQDRPKLEPFYVERYSWSQLKLLADTRKHGYMM 60
QY 61 AKAPHDFMFVKENDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNINRAAVLMLSKPILL 120
DB 61 AKAPHDFMFVKENDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNINRAAVLMLSKPILL 120
QY 121 DLFOATLDYGMYSREELLRRKRIGTVGIASVDYHOGSGTFLFQAGSGIYHVKGQPGQ 180
DB 121 DLFOATLDYGMYSREELLRRKRIGTVGIASVDYHOGSGTFLFQAGSGIYHVKGQPGQ 180
QY 181 FTQQLRPNLVETSCPNRMDDPKLCPADPDWIAFHSNDIWIINVTREERLTYVHNEL 240
DB 181 FTQQLRPNLVETSCPNRMDDPKLCPADPDWIAFHSNDIWIINVTREERLTYVHNEL 240
QY 241 ANNEEDARSAGVATFVLOEEDFDRYGVWCPKATTPSGGKILRLIYEENDESEVEIIVH 300
DB 241 ANNEEDARSAGVATFVLOEEDFDRYGVWCPKATTPSGGKILRLIYEENDESEVEIIVH 300
QY 301 TSPMLETRADSFYPKGTANPKVTFKMSIIMIDAEGRIIDVIDKELIOPPEILFEGVE 360
DB 301 TSPMLETRADSFYPKGTANPKVTFKMSIIMIDAEGRIIDVIDKELIOPPEILFEGVE 360
QY 361 YIARAGWTEGKYANSLIDRSOTELQIVLISPELFIPVEDDWMERORLIESVPDSVTEPL 420
DB 361 YIARAGWTEGKYANSLIDRSOTELQIVLISPELFIPVEDDWMERORLIESVPDSVTEPL 420
QY 421 IYVEETDIIWINHDIHFVFPQSHEEIEPIFASCEKTFGRHLKYKITSILKSKYKSSG 480
DB 421 IYVEETDIIWINHDIHFVFPQSHEEIEPIFASCEKTFGRHLKYKITSILKSKYKSSG 480
QY 481 GLPAPSDFKCPKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEBHLIYVVS 540
DB 481 GLPAPSDFKCPKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEBHLIYVVS 540
QY 541 YVNPGEVTRLDGRGSHSCCISQHCDFIFISKYSQKNPCHVCVSLYKLSPPDDPTCKTKEF 600
DB 541 YVNPGEVTRLDGRGSHSCCISQHCDFIFISKYSQKNPCHVCVSLYKLSPPDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEITFSPESITGFTLYGMLYKPHDLQPKGKYPTVLIYGGQVQL 660
DB 601 WATILDSAGPLPDYTPPEITFSPESITGFTLYGMLYKPHDLQPKGKYPTVLIYGGQVQL 660
QY 661 VNNRFGVKYFRNLTLASLYGVVVIIDNRGSHRGLKPEGAFKYMGIQIIDDQVREGLOV 720
DB 661 VNNRFGVKYFRNLTLASLYGVVVIIDNRGSHRGLKPEGAFKYMGIQIIDDQVREGLOV 720
QY 721 LASRYDFIDLDRVINGHWSYGGVLSLMALMORSIDFRVAIAGAPVTLIWIFYDTGYTERYM 780
DB 721 LASRYDFIDLDRVINGHWSYGGVLSLMALMORSIDFRVAIAGAPVTLIWIFYDTGYTERYM 780

QY 781 GHPDONEGYVGLGVAMQAEKFPSEPNRLLLHGLFDENHFAHTSILLISFLVRACKPYD 840
DB 781 GHPDONEGYVGLGVAMQAEKFPSEPNRLLLHGLFDENHFAHTSILLISFLVRACKPYD 840
QY 841 LQIYPOERHSIRVPSGEHVELHLLHYLQENLGSRIAALKVI 882
DB 841 LQIYPOERHSIRVPSGEHVELHLLHYLQENLGSRIAALKVI 882
RESULT 2
ABG61591
ID ABG61591 standard; protein; 882 AA.
AC ABG61591;
XX
DT 12-AUG-2002 (first entry)
DE Human DPPIV related serine protease DPPP-1.
XX
DE Human: serine protease; dipeptidyl peptidase IV-related protein; DPPP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinnesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX
OS Homo sapiens.
XX
PN WO200231134-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031874.
XX
PR 12-OCT-2000; 2000US-0240117P.
XX
PA (FERR) FERRING BV.
XX
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX N-PSDB; ABK83322.
XX
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
PS Claim 17; Fig 1; 113pp; English.
XX
CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypertension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinasias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPPP proteins
XX
SQ Sequence 882 AA;
Query Match 100.0%; Score 4700; DB 5; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMETQLGVPEITADCEENIESQDRPKLEPFYVERYSWSQLKLLADTRKHGYMM 60

Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTKYHYGM 60
Qy 61 AKAPHDFMFKVNDPDGPHSDRIYYLAMSNGENRENTLFYSEIPKTNRAAVLMSWKPL 120
Db 61 AKAPHDFMFKVNDPDGPHSDRIYYLAMSNGENRENTLFYSEIPKTNRAAVLMSWKPL 120
Qy 121 DLFOATLDYGMTSREELRLRRKRIGTVGIASDYHOGSGTFLFQAGSGIYHVKGDPG 180
Db 121 DLFOATLDYGMTSREELRLRRKRIGTVGIASDYHOGSGTFLFQAGSGIYHVKGDPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAIHSNDIWSNIVTREERLTYVHNE 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAIHSNDIWSNIVTREERLTYVHNE 240
Qy 241 ANNEEDARSAGVATVLOEEDRYSGYWCPCAKTTPSGGKIIRLIYEENDESEVEI 300
Db 241 ANNEEDARSAGVATVLOEEDRYSGYWCPCAKTTPSGGKIIRLIYEENDESEVEI 300
Qy 301 TSPMLETRADSFYKPTGTANPKVTFKMSIMIDABRIIDVIDKELIQPEILPEGVE 360
Db 301 TSPMLETRADSFYKPTGTANPKVTFKMSIMIDABRIIDVIDKELIQPEILPEGVE 360
Qy 361 YIARAGWPEGKYANSILLDRSOTELQIVLISPELFTPVEDDWMERORLIESVPSV 420
Db 361 YIARAGWPEGKYANSILLDRSOTELQIVLISPELFTPVEDDWMERORLIESVPSV 420
Qy 421 IYEEETDIIWNIHDIHVFPQSHHEEIEFIFASECKTGPRHLYKITSILKESKYRSG 480
Db 421 IYEEETDIIWNIHDIHVFPQSHHEEIEFIFASECKTGPRHLYKITSILKESKYRSG 480
Qy 481 GLPAPSDPKPIKEBIAITSGEWELVGRHSGNIQVDEVRRLVYFEGTKDSPLEHLYV 540
Db 481 GLPAPSDPKPIKEBIAITSGEWELVGRHSGNIQVDEVRRLVYFEGTKDSPLEHLYV 540
Qy 541 YNPGEVRLTDGRYSHSCCTIQRCDPFIISKYSNQKPNHCVLSYKLSPPEDDPTCKTE 600
Db 541 YNPGEVRLTDGRYSHSCCTIQRCDPFIISKYSNQKPNHCVLSYKLSPPEDDPTCKTE 600
Qy 601 WATILDSAGLPDYTPPEIFSPFESTGTFTLYGMLYKPHDLQPKKYPTVLFIYGGPV 660
Db 601 WATILDSAGLPDYTPPEIFSPFESTGTFTLYGMLYKPHDLQPKKYPTVLFIYGGPV 660
Qy 661 VNNRFGVKYFRLNTLASLGYVVVINDNRGSGHGLKFEAGFYKMGQIEIDQVEGLQY 720
Db 661 VNNRFGVKYFRLNTLASLGYVVVINDNRGSGHGLKFEAGFYKMGQIEIDQVEGLQY 720
Qy 721 LASRYDFIDLDRVGIHGSYGGYLSLMLMORSIDIFRVAIAGAPVTLWIFDTGYTER 780
Db 721 LASRYDFIDLDRVGIHGSYGGYLSLMLMORSIDIFRVAIAGAPVTLWIFDTGYTER 780
Qy 781 GHPDNEQGYLGSVAMQAEKFPSEPNRLLLHGLFDENVHFAHTSILLSEFLVRACKPY 840
Db 781 GHPDNEQGYLGSVAMQAEKFPSEPNRLLLHGLFDENVHFAHTSILLSEFLVRACKPY 840
Qy 841 LQIYQERHSIRVPESGHEHLLHLYLQENLGSRIAALKVI 882
Db 841 LQIYQERHSIRVPESGHEHLLHLYLQENLGSRIAALKVI 882

RESULT 3

AAG78415
ID AAG78415 standard; protein; 882 AA.
XX
AC AAG78415;
XX
DT 12-APR-2002 (first entry)
XX
DE Amino acid sequence of 21953 human prolyl oligopeptidase.
XX
KW 21953 prolyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;

KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease.
OS Homo sapiens.
XX
XX WO200179473-A2.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US040483.
XX
XX 18-APR-2000; 2000US-0197508P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers RA, Williamson M;
XX
XX WPI; 2002-034353/04.
XX
XX N-PSDB; AAH99934.
XX
XX New polypeptides 21953, member of human prolyl oligopeptidase family,
XX useful as diagnostic targets and therapeutic agents for controlling
XX cancer, lymphoma and leukemia.
XX
XX Claim 1; Page 102-103; 121pp; English.
XX
XX This invention relates to an isolated 21953 human prolyl oligopeptidase.
XX Which is cytostatic, antiarthritic, antiasthmatic, neuroprotective,
XX antithyroid, dermatological, antipsoriatic, antiasthmatic,
XX ophthalmological, antiinflammatory, nontropic, antiparkinsonian,
XX anticonvulsant, gynaecological, vasotropic, antianginal, cardiac,
XX antithrombotic, anorectic and metabolic in its action. Uses include
XX gene therapy, expression or activity of 21953 protein modulator, it is
XX useful for identifying a compound which binds to it and can be used in
XX preventing, treating or detecting a cellular proliferative or
XX differentiative disorder. The 21953 molecules can act as novel diagnostic
XX targets and therapeutic agents for controlling disorders associated with
XX the aberrant activity or degradation of peptide hormones e.g., disorders
XX associated with cell differentiation and proliferation such as cancer,
XX immune function, reproductive, neurological and cardiovascular function.
XX The 21953 molecules are thus useful for treating and preventing cellular
XX proliferative and differentiative disorders, haematopoietic neoplastic
XX disorders, immune disorders such as autoimmune diseases, diabetes
XX mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
XX neuronal disorders, demyelinating diseases, vascular disorders and
XX metabolism or pain disorders. This sequence represents the amino acid
XX sequence of 21953 human prolyl oligopeptidase
XX
XX Sequence 882 AA;

Query Match 100.0%; Score 4700; DB 5; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTKYHYGM 60
Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTKYHYGM 60
Qy 61 AKAPHDFMFKVNDPDGPHSDRIYYLAMSNGENRENTLFYSEIPKTNRAAVLMSWKPL 120
Db 61 AKAPHDFMFKVNDPDGPHSDRIYYLAMSNGENRENTLFYSEIPKTNRAAVLMSWKPL 120
Qy 121 DLFOATLDYGMTSREELRLRRKRIGTVGIASDYHOGSGTFLFQAGSGIYHVKGDPG 180
Db 121 DLFOATLDYGMTSREELRLRRKRIGTVGIASDYHOGSGTFLFQAGSGIYHVKGDPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAIHSNDIWSNIVTREERLTYVHNE 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAIHSNDIWSNIVTREERLTYVHNE 240
Qy 241 ANNEEDARSAGVATVLOEEDRYSGYWCPCAKTTPSGGKIIRLIYEENDESEVEI 300
Db 241 ANNEEDARSAGVATVLOEEDRYSGYWCPCAKTTPSGGKIIRLIYEENDESEVEI 300

QY 781 GHPDQNEQGYLGVSAMQAEKFPSEPNRLLLHGLDENVHFAHTSILLSFLVRACKPYD 840
 DB 781 GHPDQNEQGYLGVSAMQAEKFPSEPNRLLLHGLDENVHFAHTSILLSFLVRACKPYD 840
 QY 841 LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882
 DB 841 LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882

RESULT 5
 AAU74749
 ID AAU74749 standard; protein; 882 AA.
 AC AAU74749;
 XX
 XX 09-APR-2002 (first entry)
 XX
 DE Human protease PR7S-9 protein sequence.
 XX
 KW Human; protease; PR7S; gastrointestinal; Crohn's disease; cancer;
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW cell proliferative disorder; developmental disorder; epilepsy;
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
 KW reproductive disorder; endometriosis.
 KW
 XX Homo sapiens.
 OS
 XX W0200198468-A2.
 PN
 XX 27-DEC-2001.
 PD
 XX 13-JUN-2001; 2001WO-05019178.
 PF
 XX 16-JUN-2000; 2000US-0212336P.
 PR 22-JUN-2000; 2000US-0213955P.
 PR 29-JUN-2000; 2000US-0215396P.
 PR 07-JUL-2000; 2000US-0216821P.
 PR 14-JUL-2000; 2000US-0218946P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 FA
 XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
 PI Deleagane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
 PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
 PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
 PI Kallick DA;
 PI
 XX WPI: 2002-090437/12.
 DR N-PSDB; ABK12892.
 DR
 XX Twenty one human proteases (referred to as PR7S-1 to PR7S-21), useful in
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
 PT (e.g. cancer) disorders.
 XX
 XX Claim 1; Page 140-142; 177pp; English.
 ES
 XX The present invention relates to twenty one new human proteases, referred
 CC to as PR7S-1 to PR7S-21. The PR7S polynucleotides and polypeptides of the
 CC invention are useful in the diagnosis, treatment and prevention of
 CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
 CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
 CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
 CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
 CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
 CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
 CC disease and reproductive e.g. infertility and endometriosis disorders.
 CC Numerous other examples of each disorder are given in the specification.
 CC The present protein sequence represents the human protease PR7S-9 protein
 CC of the invention
 XX

SQ Sequence 882 AA;
 Query Match 100.0%; Score 4700; DB 5; Length 882;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAAAMETEOLGVEIPEFADCEENIESQDRPKLEPPYVERYSWSOLKKLLADTRKYHYMM 60
 DB 1 MAAAMETEOLGVEIPEFADCEENIESQDRPKLEPPYVERYSWSOLKKLLADTRKYHYMM 60
 QY 61 AKAPHDFMFVKRNDPDGPHSDRIYYLAMSNGENRENTFYSEIPKTIINRAAVALMSWPKLL 120
 DB 61 AKAPHDFMFVKRNDPDGPHSDRIYYLAMSNGENRENTFYSEIPKTIINRAAVALMSWPKLL 120
 QY 121 DLFOATLDYGMYSREBELLRERKRTGTGVIASDYHQSGTFLFOAGSGYHVHVGQPG 180
 DB 121 DLFOATLDYGMYSREBELLRERKRTGTGVIASDYHQSGTFLFOAGSGYHVHVGQPG 180
 QY 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFIHNSNDIWIISNIVTREERLLTYVHNL 240
 DB 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFIHNSNDIWIISNIVTREERLLTYVHNL 240
 QY 241 ANMEEDARAGVATFVLQEEFORYGYWPCAKETTPSGGKILRLIYEENDESEVEIHV 300
 DB 241 ANMEEDARAGVATFVLQEEFORYGYWPCAKETTPSGGKILRLIYEENDESEVEIHV 300
 QY 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIQPEILFEGVE 360
 DB 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIQPEILFEGVE 360
 QY 361 YIARAGWTPEGKYAWSILLDRSQTRLOIQLVISPFLFIVEDDVMERORLIESVDSVTPL 420
 DB 361 YIARAGWTPEGKYAWSILLDRSQTRLOIQLVISPFLFIVEDDVMERORLIESVDSVTPL 420
 QY 421 IIVEETTDWINIHDFHVPQSHHEEIEFIFASECKTGRHLKYITSILKESKYKSSG 480
 DB 421 IIVEETTDWINIHDFHVPQSHHEEIEFIFASECKTGRHLKYITSILKESKYKSSG 480
 QY 481 GLPAPSDFKCPKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVS 540
 DB 481 GLPAPSDFKCPKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVS 540
 QY 541 YNPGEVTRLTRDGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF 600
 DB 541 YNPGEVTRLTRDGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF 600
 QY 601 WATILDSAGPLPDYTPPEIFSESTTGTFLYGMLYKPHDLQPKKYPTVLFIYGGPOVOL 660
 DB 601 WATILDSAGPLPDYTPPEIFSESTTGTFLYGMLYKPHDLQPKKYPTVLFIYGGPOVOL 660
 QY 661 VNNRFKGVYFRNLTLASLGYYVWVIDNRGSGHRLGKEGAFKYMKGQIEIDDOVEGLQY 720
 DB 661 VNNRFKGVYFRNLTLASLGYYVWVIDNRGSGHRLGKEGAFKYMKGQIEIDDOVEGLQY 720
 QY 721 LASRYDFIDLDRVGIHWSYGGYLSLMALMORSIDIFRVAIAGAPVTLWIFYDTGYTRYM 780
 DB 721 LASRYDFIDLDRVGIHWSYGGYLSLMALMORSIDIFRVAIAGAPVTLWIFYDTGYTRYM 780
 QY 781 GHPDQNEQGYLGVSAMQAEKFPSEPNRLLLHGLDENVHFAHTSILLSFLVRACKPYD 840
 DB 781 GHPDQNEQGYLGVSAMQAEKFPSEPNRLLLHGLDENVHFAHTSILLSFLVRACKPYD 840
 QY 841 LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882
 DB 841 LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882

RESULT 6
 ADI17086
 ID ADI17086 standard; protein; 882 AA.
 XX
 AC ADI17086;
 XX

DT	15-APR-2004	(first entry)	PA	(CURA-) CURAGEN CORP.
XX			XX	Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
DE		Human NOVX protein homologue seqID 622.	PI	Li L, Gangolli EA, Fadigaru M, Anderson DW, Rastelli L, Miller CE;
XX		human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;	PI	Gerlach VL, Taupier KU, Gusev VI, Colman SD, Wolenc AR, Pena CEA;
KW		inflammation; autoimmune disorder; allergy; blood disorder;	PI	Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
KW		acquired immunodeficiency syndrome; AIDS; obesity; asthma;	XX	WPI; 2002-706998/76.
KW		immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;	XX	New NOVX polypeptides and nucleic acids, useful for preventing or
XX		Alzheimer's disease; infection; str.	PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
OS			PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX		Homo sapiens.	PT	pharmacogenomics.
PN		WO200268649-A2.	XX	Disclosure; SEQ ID NO 622; 1498pp; English.
XX		06-SEP-2002.	XX	
PD			XX	This invention relates to a novel nucleic acids, and encoded polypeptides
PF		31-JAN-2002; 2002WO-US002785.	CC	thereof, which have properties related to the stimulation of biochemical
XX			CC	or physiological responses in a cell, tissue, organ or organism.
XX		31-JAN-2001; 2001US-0265395P.	CC	Specifically, it refers to the use of biologically active fragments for
PR		31-JAN-2001; 2001US-0265412P.	CC	diagnostic and prognostic assays and furthermore in the treatment of
PR		31-JAN-2001; 2001US-0265514P.	CC	diverse pathological conditions. The present invention describes novel
PR		02-FEB-2001; 2001US-0266406P.	CC	human and murine NOVX proteins, as well as methods to modulate their
PR		05-FEB-2001; 2001US-0266467P.	CC	expression using antisense oligos, ribozymes and peptide nucleic acids.
PR		07-FEB-2001; 2001US-0266975P.	CC	The NOVX polypeptides, polynucleotides and antibodies are useful in
PR		07-FEB-2001; 2001US-0267057P.	CC	treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
PR		08-FEB-2001; 2001US-0267459P.	CC	atherosclerosis, cancer and diabetes. Furthermore, they may be used in
PR		09-FEB-2001; 2001US-0267823P.	CC	treating or preventing diseases such as inflammation, autoimmune
PR		15-FEB-2001; 2001US-0271664P.	CC	disorders, allergies, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
PR		26-FEB-2001; 2001US-0271839P.	CC	(AIDS), obesity, Alzheimer's disease, infections, stroke, muscular dystrophy
PR		27-FEB-2001; 2001US-0271855P.	CC	arthritis, anti-HIV, antidiabetic, immunosuppressive, antiallergic,
PR		02-MAR-2001; 2001US-0273046P.	CC	cytostatic, cardiant, antiinflammatory, antihypertensive, anorectic,
PR		02-MAR-2001; 2001US-0275925P.	CC	haemostatic, neurotropic, antibacterial, virucide, antiparasitic,
PR		14-MAR-2001; 2001US-0275947P.	CC	neuroprotective, nontoxic, antidiabetic, antihypertensive, anorectic,
PR		14-MAR-2001; 2001US-0275950P.	CC	relaxant and anticonvulsant. In addition, they are useful in screening
PR		14-MAR-2001; 2001US-0275980P.	CC	assays to identify small molecules that modulate or inhibit, for example,
PR		15-MAR-2001; 2001US-0276448P.	CC	neurogenesis, wound healing and angiogenesis. The nucleic acids are also
PR		15-MAR-2001; 2001US-0276450P.	CC	used as in chromosome mapping, tissue typing, preventive medicine and
PR		16-MAR-2001; 2001US-0276597P.	CC	pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
PR		16-MAR-2001; 2001US-0276598P.	CC	of the invention.
PR		20-MAR-2001; 2001US-0278652P.	XX	Sequence 882 AA;
PR		26-MAR-2001; 2001US-0278775P.	XX	Query Match 100.0%; Score 4700; DB 5; Length 882;
PR		26-MAR-2001; 2001US-0278778P.	XX	Best Local Similarity 100.0%; Pred. No. 0;
PR		29-MAR-2001; 2001US-0279884P.	XX	Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
PR		30-MAR-2001; 2001US-0280147P.	XX	
PR		11-APR-2001; 2001US-0282992P.	Qy	1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
PR		11-APR-2001; 2001US-0283083P.	Db	1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
PR		20-APR-2001; 2001US-0285133P.	Qy	61 AKAPHDFMFVKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAAVLMSWKPLL 120
PR		23-APR-2001; 2001US-0285749P.	Db	61 AKAPHDFMFVKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAAVLMSWKPLL 120
PR		03-MAY-2001; 2001US-0288327P.	Qy	121 DLFOATLDYGMYSREELRLERKRIGTVGIASYDHQSGTFLFQAGSGYHVHKGDPQG 180
PR		03-MAY-2001; 2001US-0288504P.	Db	121 DLFOATLDYGMYSREELRLERKRIGTVGIASYDHQSGTFLFQAGSGYHVHKGDPQG 180
PR		29-MAY-2001; 2001US-0294047P.	Qy	121 DLFOATLDYGMYSREELRLERKRIGTVGIASYDHQSGTFLFQAGSGYHVHKGDPQG 180
PR		30-MAY-2001; 2001US-0294473P.	Db	121 DLFOATLDYGMYSREELRLERKRIGTVGIASYDHQSGTFLFQAGSGYHVHKGDPQG 180
PR		08-JUN-2001; 2001US-0296964P.	Qy	181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFHNSNDIWSNIVTFRERLLTVHNEL 240
PR		18-JUN-2001; 2001US-0298959P.	Db	181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFHNSNDIWSNIVTFRERLLTVHNEL 240
PR		19-JUN-2001; 2001US-0299324P.	Qy	241 ANMEEDARSAGVATFVLQEEFDRYSGYWCWPKAETTPSGGKILRLIENDESEVEIHV 300
PR		13-AUG-2001; 2001US-0312020P.	Db	241 ANMEEDARSAGVATFVLQEEFDRYSGYWCWPKAETTPSGGKILRLIENDESEVEIHV 300
PR		16-AUG-2001; 2001US-0312889P.	Qy	301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIQPFELFEGVE 360
PR		16-AUG-2001; 2001US-0312908P.	Db	301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIQPFELFEGVE 360
PR		21-AUG-2001; 2001US-0313390P.	Qy	
PR		28-AUG-2001; 2001US-0315470P.	Db	
PR		31-AUG-2001; 2001US-0316447P.	Qy	
PR		07-SEP-2001; 2001US-0318115P.	Db	
PR		07-SEP-2001; 2001US-0318118P.	Qy	
PR		12-SEP-2001; 2001US-0318740P.	Db	
PR		19-SEP-2001; 2001US-0323379P.	Qy	
PR		18-OCT-2001; 2001US-0330245P.	Db	
PR		18-OCT-2001; 2001US-0330308P.	Qy	
PR		14-NOV-2001; 2001US-0332701P.	Db	
XX			XX	

QY 361 YIARAGWTPGKYAWMSILLDRSQTRLOIVLISPELFIPIVEDDWMERQRLIESVPDSTVPL 420
 DB |||||
 QY 361 YIARAGWTPGKYAWMSILLDRSQTRLOIVLISPELFIPIVEDDWMERQRLIESVPDSTVPL 420
 DB |||||
 QY 421 IYYEETDIWINIHDPVPPQSHHEIEFIFASECKTGFRHLKYITSILKESKYKSSG 480
 DB |||||
 QY 421 IYYEETDIWINIHDPVPPQSHHEIEFIFASECKTGFRHLKYITSILKESKYKSSG 480
 DB |||||
 QY 481 GLPAPDFKPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLVVS 540
 DB |||||
 QY 481 GLPAPDFKPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLVVS 540
 DB |||||
 QY 541 YVNPGEVTRLTDRGYSHSCCISQCHDFFIKYSKNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
 DB |||||
 QY 541 YVNPGEVTRLTDRGYSHSCCISQCHDFFIKYSKNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
 DB |||||
 QY 601 WATILDSAGPLDPTPEIFPFSTTGFTLYGMLYKPHDLQPKKYPTVLFYIGGPOVQL 660
 DB |||||
 QY 601 WATILDSAGPLDPTPEIFPFSTTGFTLYGMLYKPHDLQPKKYPTVLFYIGGPOVQL 660
 DB |||||
 QY 661 VNNRFGVKYFRNLNTLASLYGVVVVINDNRGSHRGGLKPEGAPKYKMGQIEIDDDQVEGLQY 720
 DB |||||
 QY 661 VNNRFGVKYFRNLNTLASLYGVVVVINDNRGSHRGGLKPEGAPKYKMGQIEIDDDQVEGLQY 720
 DB |||||
 QY 721 LASRYDFIDLRVGIHWSYGGYLSLMLMQRSDIFRVAIAGAPVTLWIFDYGTYERYM 780
 DB |||||
 QY 721 LASRYDFIDLRVGIHWSYGGYLSLMLMQRSDIFRVAIAGAPVTLWIFDYGTYERYM 780
 DB |||||
 QY 781 GHPODQGGYVLSVAMQAEKFPSEPNRLLHGFLENVHFAHTSILLSLVLRAGKPYD 840
 DB |||||
 QY 781 GHPODQGGYVLSVAMQAEKFPSEPNRLLHGFLENVHFAHTSILLSLVLRAGKPYD 840
 DB |||||
 QY 841 LOIYQPRHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882
 DB |||||
 QY 841 LOIYQPRHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882
 DB |||||

RESULT 7

ABU07720

ID ABU07720 standard; protein; 882 AA.

XX AC

XX ABU07720;

XX DT

XX 19-MAY-2003 (first entry)

XX DE

XX Human serine protease HIPHUM46.

XX KW

Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;
 serine protease activity modulation; dipeptidyl peptidase activity;
 musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
 Alzheimer's disease; parasupranuclear palsy; Huntington's disease;
 amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
 irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;
 haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
 colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
 multiple sclerosis.

XX OS

XX Homo sapiens.

XX FH

XX Key Location/Qualifiers

XX FT

XX Region 259..260

XX FT

XX Active-site /note= "Paired glutamates of the beta propeller domain"

XX FT

XX Active-site /label= Catalytic_serine_residue

XX FT

XX Active-site /label= Catalytic_aspartate_residue

XX FT

XX Active-site /label= Catalytic_histidine_residue

XX XX

XX GB2374869-A.

XX PN

XX 30-OCT-2002.

XX PD

XX XX

PF 22-JAN-2002; 2002GB-00001404.

XX 23-JAN-2001; 2001GB-00001760.

XX (GLAX) GLAXO GROUP LTD.

XX Edbrooke MR, Lewis AP;

XX WPI; 2003-150703/15.

XX N-PSDB; ABX12255.

XX Identifying modulators of serine protease activity useful for treating

XX musculoskeletal diseases, by contacting cell expressing a novel serine

XX protease polypeptide with a compound and monitoring serine protease

XX activity.

XX Claim 10; Page 26-29; 38pp; English.

XX The invention relates to a method of identifying a substance that

XX modulates serine protease activity, comprising contacting a cell such as

XX a neuronal cell, lung cell, intestinal cell or a cell infected with a

XX virus, expressing a serine protease polypeptide (HIPHUM 46), or its

XX variant having dipeptidyl peptidase activity, or a serine protease

XX isolated from the cell with a test substance and monitoring for serine

XX protease activity. The method is useful for identifying a substance that

XX modulates serine protease activity. A modulator of the serine protease is

XX useful in the manufacture of a medicament for treatment or prophylaxis of

XX a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus

XX infection, Alzheimer's disease, parasupranuclear palsy, myotonic

XX dystrophy, Huntington's disease or amyotrophic lateral sclerosis.

XX Additional disease that may be treated using modulators of the serine

XX protease include malabsorption syndromes, irritable bowel syndrome, lung

XX disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,

XX rectal polyps, small bowel tumours, colorectal tumours, anaemia,

XX dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple

XX sclerosis. The present sequence represents the amino acid sequence of the

XX human serine protease HIPHUM46

XX SQ

XX Sequence 882 AA;

Query Match

Best Local Similarity 100.0%; Score 4700; DB 6; Length 882;

Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAMETEQLGVEIFETADCEENIESQDRPKLEFPFVVERYSWSQLKLADTRKYHYGM 60

DB 1 MAAMETEQLGVEIFETADCEENIESQDRPKLEFPFVVERYSWSQLKLADTRKYHYGM 60

QY 61 AKAPDHFMVKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAAAMLWSKPL 120

DB 61 AKAPDHFMVKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAAAMLWSKPL 120

QY 121 DLFOATLDYGMYSREBELLRERKRICTVGIASYDYHQSGTFLFOAGSGIYHVKGQPG 180

DB 121 DLFOATLDYGMYSREBELLRERKRICTVGIASYDYHQSGTFLFOAGSGIYHVKGQPG 180

QY 181 FTQQLPLNVLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLLTYVHNEL 240

DB 181 FTQQLPLNVLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLLTYVHNEL 240

QY 241 ANMEEDARSAGVATFVLQBEFDYSGYWWCPKAETTPSGGKILRLIYENDESEVEIIHV 300

DB 241 ANMEEDARSAGVATFVLQBEFDYSGYWWCPKAETTPSGGKILRLIYENDESEVEIIHV 300

QY 301 TSPMLETRRADSFYRPTGTANPKVTFKMSIEMDAEGRIIDVIDKELIQFEILFEGVE 360

DB 301 TSPMLETRRADSFYRPTGTANPKVTFKMSIEMDAEGRIIDVIDKELIQFEILFEGVE 360

QY 361 YIARAGWTPGKYAWMSILLDRSQTRLOIVLISPELFIPIVEDDWMERQRLIESVPDSTVPL 420

DB 361 YIARAGWTPGKYAWMSILLDRSQTRLOIVLISPELFIPIVEDDWMERQRLIESVPDSTVPL 420

QY 421 IYYEETDIWINIHDPVPPQSHHEIEFIFASECKTGFRHLKYITSILKESKYKSSG 480

Tue Apr 18 08:18:34 2006

u8-10-825-632-1.rag

421	DB	LIYEVTTDWINIHDIPIHVFPQSHHEEIEIFASECKTGFRHLYKITSILKESKYKRSG	480
481	QY	GLPAPSDFKPIKEEIAITTSGEWVLGRHGSNIQVDEVRLLYVFEGTKDSPLEHLYVVS	540
481	DB	GLPAPSDFKPIKEEIAITTSGEWVLGRHGSNIQVDEVRLLYVFEGTKDSPLEHLYVVS	540
541	QY	VYNPGEVTRLTDRGYSHSCCCI SQHCDFFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF	600
541	DB	VYNPGEVTRLTDRGYSHSCCCI SQHCDFFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF	600
601	QY	WATILDSAGPLPDYTPPEIIPFSESTGFTLYGMLYKPHDLOPKKYPTVLFTYGGPQVL	660
601	DB	WATILDSAGPLPDYTPPEIIPFSESTGFTLYGMLYKPHDLOPKKYPTVLFTYGGPQVL	660
661	QY	VNNRFGVKYFRNLNTLASLGVVVVVVIDNRGSHRGCLKFEGAFKYMKGQTEIDDDVEGLQY	720
661	DB	VNNRFGVKYFRNLNTLASLGVVVVVVIDNRGSHRGCLKFEGAFKYMKGQTEIDDDVEGLQY	720
721	QY	LASRYDFIDLDRVGTHGWSYGGYISLMALMQRSDFRVAITAGAPVTLWFTYGTGYTERVM	780
721	DB	LASRYDFIDLDRVGTHGWSYGGYISLMALMQRSDFRVAITAGAPVTLWFTYGTGYTERVM	780
781	QY	GHPDQNEQGYILGVSAMQAQEPSPFNRLLLHGFLDENVHFAHTSILLSFLVRACKPYD	840
781	DB	GHPDQNEQGYILGVSAMQAQEPSPFNRLLLHGFLDENVHFAHTSILLSFLVRACKPYD	840
841	QY	LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIALVKVI	882
841	DB	LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIALVKVI	882

RESULT 8

ADI17085
ID ADI17085 standard: protein: 883 AA.

ID ADI17085
XX
AC ADI17085;

XX 15-APR-2004 (first entry)

XX Murine NOVX protein homologue SegID 621.

XX mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (IgA nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infect.

KW
 XX
 OS

XX
PN WO200268649-A2.

XX
PD 06-SEP-2002.

31 -JAN-2002: 2002WO-US002785.

31 -JAN-2001. 2001US-0265395P.

PR 31-JAN-2001; 2001US-0265514P
PR 31-JAN-2001; 2001US-0265412P
PR 31-JAN-2001; 2001US-0265514P

PR 31-JAN-2001; 2001US-0265314F
PR 31-JAN-2001; 2001US-0265517P
PR 31-JAN-2001; 2001US-0265405D

PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.

PR 07-FEB-2001; 2001US-0266975P.
PR 07-FEB-2001; 2001US-0267057P.

PR 08-FEB-2001; 2001US-0267459P.
PR 09-FEB-2001; 2001US-0267823P.

PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.

PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.

PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.

PR 14-MAR-2001; 2001US-0275925P.

PR	14-MAR-2001;	2001US-0275947P
PR	14-MAR-2001;	2001US-0275950P
PR	14-MAR-2001;	2001US-0275989P
PR	15-MAR-2001;	2001US-0276448P
PR	15-MAR-2001;	2001US-0276450P
PR	15-MAR-2001;	2001US-0276459P
PR	16-MAR-2001;	2001US-0276439P
PR	16-MAR-2001;	2001US-0276769P
PR	20-MAR-2001;	2001US-0278652P
PR	26-MAR-2001;	2001US-0278775P
PR	26-MAR-2001;	2001US-0278788P
PR	29-MAR-2001;	2001US-0279884P
PR	29-MAR-2001;	2001US-0279884P
PR	30-MAR-2001;	2001US-0280147P
PR	11-APR-2001;	2001US-0283992P
PR	11-APR-2001;	2001US-0283983P
PR	20-APR-2001;	2001US-0285133P
PR	20-APR-2001;	2001US-0285749P
PR	03-MAY-2001;	2001US-0286327P
PR	03-MAY-2001;	2001US-0288504P
PR	30-MAY-2001;	2001US-0294047P
PR	30-MAY-2001;	2001US-0294473P
PR	08-JUN-2001;	2001US-0296964P
PR	18-JUN-2001;	2001US-0296959P
PR	19-JUN-2001;	2001US-0299324P
PR	13-AUG-2001;	2001US-0312020P
PR	16-AUG-2001;	2001US-0312889P
PR	16-AUG-2001;	2001US-0312908P
PR	21-AUG-2001;	2001US-0313390P
PR	28-AUG-2001;	2001US-0315470P
PR	31-AUG-2001;	2001US-0316447P
PR	07-SEP-2001;	2001US-0318115P
PR	12-SEP-2001;	2001US-0318118P
PR	17-SEP-2001;	2001US-0318740P
PR	18-SEP-2001;	2001US-0323379P
PR	19-OCT-2001;	2001US-0323024P
PR	18-OCT-2001;	2001US-0330308P
PR	14-NOV-2001;	2001US-0332701P

(CURA-) CURAGEN CORP.

Thernev VT, Spytek KA, Zernhusen BD, Patturajan M, Shimkets RA;
Li L, Gangoli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
Gerlach V, Taupier RJ, Gusev VV, Colman SD, Wolenc AR, Pena CRA;
Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Burgess CE;
WPI: 2002-706998/76.

New NOVX polypeptides and nucleic acids, useful for preventing or treating NOVX-associated disorders, e.g. cancer, cardiomyopathy, atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or pharmacogenomics.

Disclosure: SEO ID NO 621; 1498pp; English.

This invention relates to a novel nucleic acids, and encoded polypeptides thereof which have properties related to the stimulation of biochemical or physiological responses in a cell, tissue, organ or organism. Specifically, it refers to the use of biologically active fragments for diagnostic and prognostic assays and furthermore in the treatment of diverse pathological conditions. The present invention describes novel human and murine NOVX proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOVX polypeptides, polynucleotides and antibodies are useful in treating or preventing NOVX-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including: cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic, antiasthmatic, neurotrophic, antiarthritic, hepatotropic.

CC neuroprotective, neurotropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.

Sequence 883 AA:

Query Match	96.4%;	Score 4528.5;	DB 5;	Length 883;
Best Local Similarity	95.7%;	Pred. No. 0;		
Matches	845;	Conservative 18;	Mismatches 19;	Indels 1;
Gaps	1;			
Qy	1	MAAAETEOLGVEIETADCEE-NIESQDRPKLEPFVVERYSWSQKLKLADTRKYHGYM	59	
Db	1	MAAAETEOLGVEIETADCEE-NIESQDRPKLEPFVVERYSWSQKLKLADTRKYHGYM	60	
Qy	60	MAKAPHDFMFVKRNDPDGPHSDRIYVYIAMSSENRENTLFYSEIPTKINRAAVLMSWPKL	119	
Db	61	MAKAPHDFMFVKRNDPDGPHSDRIYVYIAMSSENRENTLFYSEIPTKINRAAVLMSWPKL	120	
Qy	120	LDLFQATLDYGMYSREBELLRKRKIGTVGTIASYDHYHQSQTPLFOAGSGIYHVKGQGP	179	
Db	121	LDLFQATLDYGMYSREBELLRKRKIGTVGTIAAYDHYHQSQTPLFOAGSGIYHDKGGPH	180	
Qy	180	GFTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWLSTNVTVREERLTYVNE	233	
Db	181	GFTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWLSTNVTVREERLTYVNE	240	
Qy	240	LANNEDARSAGVAIFVLQEBFDRYSGYWCPCAEITPSSGKILRLIYEENDESEVEIHH	299	
Db	241	LANNEDPRSAGVAIFVLQEBFDRYSGYWCPCAEITPSSGKILRLIYEENDESEVEIHH	300	
Qy	300	VTSPLMETRRASFRYPKGTGTANPKVTPKMGSEIMIDAEGRIDVIDKELIQPEILPFGV	359	
Db	301	VTSPLMETRRASFRYPKGTGTANPKVTPKMGSEIVVDAAGGIDVIDKELIQPEILPFGV	360	
Qy	360	EYARAGWTPEGKYAWSILLDRSQTRQLQIVLISPELPIPVEDDWMERQRLIESVPDSTVP	419	
Db	361	EYARAGWTPEGKHAWSILLDRSQTHQLQIVLISPELPIPVEDDWMERQRLIESVPDSTVP	420	
Qy	420	LIIVYEETDIWINIHDIHVFPQSHHEEIEIFIPASECKTGFRHLKYITSLKESKYGSS	479	
Db	421	LIIVYEETDIWINIHDIHVFPQTHEDIEIFIPASECKTGFRHLKYITSLKESKYGSS	480	
Qy	480	GGLPAPSDFKCPKEEIAITSGEWEVLGRHGSNTQDVDEVRRLVYFGTWDSPLEHLHYV	539	
Db	481	GGLPAPSDFKCPKEEITITSGEWEVLGRHGSNTWDEARKLVYFEGTKDPSLEHLHYT	540	
Qy	540	SYNVPGEVRLTRDRGYSHSCCISQCHDFFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKE	599	
Db	541	SYANPGEVRLTRDRGYSHSCCISLRCDFFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKE	600	
Qy	600	FWATILDSAGPLPDYTPPEISFSESTTGFTLYGMLYKPHDLQPKKYPTVLFIYGGPQVQ	659	
Db	601	FWATILDSAGPLPDYTPPEISFSESTTGFTLYGMLYKPHDLQPKKYPTVLFIYGGPQVQ	660	
Qy	660	LNNRFGVKYKPLANTLASLGIVVVVVVDNRGSCHRGLKFEGAFKXMGQIIEDDQVEGLQ	719	
Db	661	LNNRFGVKYKPLANTLASLGIVVVVVVDNRGSCHRGLKFEGAFKXMGQIIEDDQVEGLQ	720	
Qy	720	YLASRVDFDLDVRVGIHGSYGGYLSIMALWQRSDIFRVAITAGAPVTLMIFYDTGTYERY	779	
Db	721	YLASQVDFDLDVRVGIHGSYGGYLSIMALWQRSDIFRVAITAGAPVTLMIFYDTGTYERY	780	
Qy	780	MGPDPQNEQGYVILGVSVAQAEKFPSEPNRLLLHGFLDENVVHAFTSILLSFVLRAKGPY	839	
Db	781	MGPDPQNEQGYVILGVSVAQAEKFPSEPNRLLLHGFLDENVVHAFTSILLSFVLRAKGPY	840	
Qy	840	DLQIYPOERHSIRVPESGHEYLHLLHYLOENLGSRTAAALKVI	882	
Db	841	DLQIYPOERHSIRVPESGHEYLHLLHYLOENLGSRTAAALKVI	883	

Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYGYMM 60
Qy 61 AKAPHDFMFKRNDPDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
Db 61 AKAPHDFMFKRNDPDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
Qy 121 DLFOATLDYGMYSREBELLRERKRIGTVGSIASYDHQSGTFLFOAGSGIYHVKGDPQG 180
Db 121 DLFO----- 124
Qy 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYVHNEL 240
Db 125 --OQPLRPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYVHNEL 182
Qy 241 ANMEEDARSAGVATFVLQEFDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
Db 183 ANMEEDARSAGVATFVLQEFDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 242
Qy 301 TSPMLETRRADSPRYPKTGTPANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db 243 TSPMLETRRADSPRYPKTGTPANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 302
Qy 361 YIARAGWTPGKYAWSILLDRSQTRLIQVILISPELFIPEVDVMMERQRLIESVPDSVTPL 420
Db 303 YIARAGWTPGKYAWSILLDRSQTRLIQVILISPELFIPEVDVMMERQRLIESVPDSVTPL 362
Qy 421 ILYEETTDIWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYITSLILKESKYRSSG 480
Db 363 ILYEETTDIWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYITSLILKESKYRSSG 422
Qy 481 GLPAPSDFKPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYPEGTKDSPLEHLYVVS 540
Db 423 GLPAPSDFKPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYPEGTKDSPLEHLYVVS 482
Qy 541 YNPGEVTRLTDRGYSHSCCISQHCDFPISKYNOXNPHCVSLYKLSPEDDPTCKTKEF 600
Db 483 YNPGEVTRLTDRGYSHSCCISQHCDFPISKYNOXNPHCVSLYKLSPEDDPTCKTKEF 542
Qy 601 WATILSAGPLDPTPEIFSPFESTTGFTLYGMLYKPHDLOPKKYPTVLFYGGQVQL 660
Db 543 WATILSAGPLDPTPEIFSPFESTTGFTLYGMLYKPHDLOPKKYPTVLFYGGQVQL 602
Qy 661 VNNRFKGVKFRNLTLASGYVWVVDNRGSGHRLKPEGAFKYMGIIEIDQVEGLQY 720
Db 603 VNNRFKGVKFRNLTLASGYVWVVDNRGSGHRLKPEGAFKYMGIIEIDQVEGLQY 662
Qy 721 LASRYDFIDLDRVGIHWSYGYLSLMLMQRSDIFRVAIAGAPVTLWIFDGTGYTERYM 780
Db 663 LASRYDFIDLDRVGIHWSYGYLSLMLMQRSDIFRVAIAGAPVTLWIFDGTGYTERYM 722
Qy 781 GHPDNEQGYLGSVAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSPFLVRAGKPYD 840
Db 723 GHPDNEQGYLGSVAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSPFLVRAGKPYD 782
Qy 841 LOIYQBRHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882
Db 783 LOIYQBRHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 824

RESULT 11

ABB97361 ID ABB97361 standard; protein; 782 AA.
XX AC ABB97361;
XX DT 27-JUN-2002 (first entry)
XX XX Novel human protein SEQ ID NO: 629.
XX Human; antianaemic; vulnary; antinflammatory; immunomodulator;
KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; BST;
KW expressed sequence tag.

XX OS Homo sapiens.
XX WO200222660-A2.
XX 21-MAR-2002.
XX 10-SEP-2001; 2001WO-US026015.
XX 11-SEP-2000; 2000US-00659671.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-292408/33.
XX N-PSDB; ABB32547.
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
XX Example 2; SEQ ID NO 629; 509pp; English.
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention
XX Sequence 782 AA;

Query Match 87.6%; Score 4118; DB 5; Length 782;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 100; Gaps 1;
Qy 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYGYMM 60
Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYGYMM 60
Qy 61 AKAPHDFMFKRNDPDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
Db 61 AKAPHDFMFKRNDPDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
Qy 121 DLFOATLDYGMYSREBELLRERKRIGTVGSIASYDHQSGTFLFOAGSGIYHVKGDPQG 180
Db 121 DLFOATLDYGMYSREBELLRERKRIGTVGSIASYDHQSGTFLFOAGSGIYHVKGDPQG 180
Qy 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYVHNEL 240
Db 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYVHNEL 240
Qy 241 ANMEEDARSAGVATFVLQEFDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
Db 241 ANMEEDARSAGVATFVLQEFDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
Qy 301 TSPMLETRRADSPRYPKTGTPANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRRADSPRYPKTGTPANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
Qy 361 YIARAGWTPGKYAWSILLDRSQTRLIQVILISPELFIPEVDVMMERQRLIESVPDSVTPL 420
Db 361 YIARAGWTPGKYAWSILLDRSQTRLIQVILISPELFIPEVDVMMERQRLIESVPDSVTPL 420
Qy 421 ILYEETTDIWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYITSLILKESKYRSSG 480
Db 421 ILYEETTDIWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYITSLILKESKYRSSG 480


```

Db 707 ----- 706
Qy 781 GHPDQEQYGLGVAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSFLVRACKPYD 840
Db 707 ----- 706
Qy 841 LQIYPERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882
Db 707 --IYPOERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 746

RESULT 13
ADT04044
ID ADT04044 standard; protein; 738 AA.
XX
AC ADT04044;
XX
DT 30-DEC-2004 (first entry)
XX
DE Human protein modification and maintenance molecule protein SeqID24.
XX
DE protein modification and maintenance molecule; PMMM; cytosstatic;
KW immunomodulator; expression; immune disorder; cancer; human.
XX
OS Homo sapiens.
XX
PN WO2004084806-A2.
XX
PD 07-OCT-2004.
XX
PF 16-MAR-2004; 2004WO-US008006.
XX
PR 21-MAR-2003; 2003US-0456864P.
XX
PR 03-APR-2003; 2003US-0460512P.
XX
PR 19-MAY-2003; 2003US-0472027P.
XX
PR 30-MAY-2003; 2003US-0475072P.
XX
PR 30-MAY-2003; 2003US-0475190P.
XX
PR 24-JUN-2003; 2003US-0482575P.
XX
PR 13-AUG-2003; 2003US-0495151P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Baughn MR, Marquis JP, Kable AE, Chawla NK, Emerling BM, Lee SY;
PI Hafalia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura LL;
PI Yang J, Jin P, Becha SD, Wilson AD;
XX
DR WPI; 2004-710256/69.
DR N-PSDB; ADT04072.
XX
PT New human protein modification and maintenance molecules (PMMM)
PT polypeptide, useful in preparing a composition for treating a disease
PT associated with decreased expression or overexpression of PMMM e.g.,
PT cancer.
XX
PS Claim 1; SEQ ID NO 24; 222pp; English.
XX
CC This invention relates to novel human protein modification and
CC maintenance molecule (PMMM) proteins and the DNA sequences which encode
CC them. The invention may be useful for the production of compounds with a
CC cytosstatic or immunomodulator activity. The proteins are useful in
CC preparing a composition for diagnosing or treating a disease or condition
CC associated with decreased expression or overexpression of functional
CC PMMM, for example immune disorders or cancer. The present sequence is
CC that of a human protein modification and maintenance molecule (PMMM)
CC protein of the invention.
XX
SQ Sequence 738 AA;

Query Match 81.2%; Score 3817; DB 8; Length 738;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 732; Conservative 1; Mismatches 5; Indels 144; Gaps 3;
Qy 1 MAAMETEQLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQLKLLADTRKHGYMM 60

```

```

Db 1 MAAMETEQLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQLKLLADTRKHGYMM 60
Qy 61 AKAPHDFMFKRNDPDGPHSDRIYYLAWSGENRENTLFYSEIPKTIINRAAIVLMSWPKLL 120
Db 61 AKAPHDFMFKRNDPDGPHSDRIYYLAWSGENRENTLFYSEIPKTIINRAAIVLMSWPKLL 120
Qy 121 DLFOATLDYGMYSREBELLRERKRTGTGVIASDYHQSGCTFLQAGSGIVHVKGQPG 180
Db 121 DLFOATLDYGMYSREBELLRERKRTGTGVIASDYHQSGCTFLQAGSGIVHVKGQPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWTFATHSNDIWI SINIVTREERLLYYVNNEL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWTFATHSNDIWI SINIVTREERLLYYVNNEL 240
Qy 241 ANMEEDARSAGVATFVLOEEFDRYSGYMWCPKAEITTPSGGKILRLIYEENDESEVEIHV 300
Db 241 ANMEEDARSAGVATFVLOEEFDRYSGYMWCPKAEITTPSGGKILRLIYEENDESEVEIHV 300
Qy 301 TSPMLETRRADSPRYPKTGTANPKVTFKMEISEIMIDAGRIIDVIDKELIQFELPFGVE 360
Db 277 -----TANPKVTFKMEISEIMIDAGRIIDVIDKELIQFELPFGVE 317
Qy 361 YIARAGWTEGKYAMSIILLDRSQTRLQIVLISPFLIPVEDDVMERQRLIESVDSVTPL 420
Db 318 YIARAGWTEGKYAMSIILLDRSQTRLQIVLISPFLIPVEDDVMERQRLIESVDSVTPL 377
Qy 421 IYYEETTDIWINIHDFHVFPQSHBEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480
Db 378 IYYEETTDIWINIHDFHVFPQSHBEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 437
Qy 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
Db 438 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 497
Qy 541 YVNPGEVTRLTDRGYSHSCCISQHCDFIFISKYSNOKNPHCVSLYKLSPEDDPTCKTEF 600
Db 498 YVNPGEVTRLTDRGYSHSCCISQHCDFIFISKYSNOKNPHCVSLYKLSPEDDPTCKTEF 556
Qy 601 WATILDSAGLPDYPTEPFISFESTTGTLYGMLYKPHDLQPGKKYPTVFLFYGGPQVOL 660
Db 557 WATILDSAGLPDYPTEPFISFESTTGTLYGMLYKPHDLQPGKKYPTVFLFYGGPQVOL 613
Qy 661 VNNRFKGVKYFRLNTLASLGYVWVIDNRGSCHRGLKPEGAFKYMKGQIEIDDDQVEGLQY 720
Db 614 ----- 613
Qy 721 LASRYDFIDLDRVGIHGSYGGVLSLMLMQRSDIFRVAIAGAPVTLWIFDYDTGYTERYM 780
Db 614 -----VAIAGAPVTLWIFDYDTGYTERYM 636
Qy 781 GHPDQEQYGLGVAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSFLVRACKPYD 840
Db 637 GHPDQEQYGLGVAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSFLVRACKPYD 696
Qy 841 LQIYPERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882
Db 697 LQIYPERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 738

RESULT 14
ABB97362
ID ABB97362 standard; protein; 724 AA.
XX
AC ABB97362;
XX
DT 27-JUN-2002 (first entry)
XX
DE Novel human protein SEQ ID NO: 630.
XX
KW Human; antianaemic; vulnerary; antinflammatory; immunomodulator;
KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;

```


CC (DPRP) . The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPRP proteins
XX
XX
SQ Sequence 658 AA;

Query Match 74.6%; Score 3504; DB 5; Length 658;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAMETEQLGVEIFETADCEINTESQDRPKLEPFYVERYSWSQLKLLADTRKYHGMM 60
Db 1 MAAMETEQLGVEIFETADCEINTESQDRPKLEPFYVERYSWSQLKLLADTRKYHGMM 60

QY 61 AKAPHDFMFKENDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNRAAVLMLSWKPLL 120
Db 61 AKAPHDFMFKENDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNRAAVLMLSWKPLL 120

QY 121 DLFOATLDYGMYSREELLRRERKIGTVGIASDYHOGSGTFLFQAGSGIYHVKGQPG 180
Db 121 DLFOATLDYGMYSREELLRRERKIGTVGIASDYHOGSGTFLFQAGSGIYHVKGQPG 180

QY 181 FTQOPLRNLTSCPNRMDPKLCPADPDWIAFIHNDIWI SNIVTREERLTYVHNL 240
Db 181 FTQOPLRNLTSCPNRMDPKLCPADPDWIAFIHNDIWI SNIVTREERLTYVHNL 240

QY 241 ANNEEDARSAGVATVLOEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
Db 241 ANNEEDARSAGVATVLOEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300

QY 301 TSPMLETRRADSFYPKTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPPEILFEGVE 360
Db 301 TSPMLETRRADSFYPKTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPPEILFEGVE 360

QY 361 YIARAGWTEGKYANSILLDRSQTRKQLVILISPELFPVEDDVMERORLIESVPDSVTPL 420
Db 361 YIARAGWTEGKYANSILLDRSQTRKQLVILISPELFPVEDDVMERORLIESVPDSVTPL 420

QY 421 IYVEETTDIWINIHDIHFVFPQSHEEIEFIFASECKTGFPHLYKITSILKESKYKSSG 480
Db 421 IYVEETTDIWINIHDIHFVFPQSHEEIEFIFASECKTGFPHLYKITSILKESKYKSSG 480

QY 481 GLPAPSDFKCIKEBIATSGWEVLGRHGSNIQDEVRRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDFKCIKEBIATSGWEVLGRHGSNIQDEVRRLVYFEGTKDSPLEHLYVVS 540

QY 541 YNPGSVTLTRDGRYSHSCCISQCHDFIISKYSNOKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db 541 YNPGSVTLTRDGRYSHSCCISQCHDFIISKYSNOKNPHCVSLYKLSPPEDDPTCKTKEF 600

QY 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFIYGG 655
Db 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFIYGG 655

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 12:46:23 ; Search time 46 Seconds
(without alignments)
1844.852 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAAMTEQLGVIEIFETADC.....HLHYLOENLGSRIAALKVI 882

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*
1: Piri.*
2: Piri2.*
3: Piri3.*
4: Piri4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	922.5	19.6	931	T32919	hypothetical prote
2	842	17.9	738	A87516	dipeptidyl peptida
3	754	16.0	741	JC5142	X-Pro dipeptidyl-p
4	620	13.2	711	S66261	X-Pro dipeptidyl-p
5	529	11.3	766	CDHU26	dipeptidyl-peptida
6	519.5	11.1	793	T41703	dipeptidyl aminope
7	518.5	11.0	760	S23752	dipeptidyl-peptida
8	510.5	10.9	792	A39914	dipeptidyl-peptida
9	484.5	10.3	818	A30107	dipeptidyl aminope
10	465	9.9	931	A49737	dipeptidyl aminope
11	451.5	9.6	799	T25174	hypothetical prote
12	445.5	9.5	779	T25173	dipeptidyl aminope
13	440	9.4	803	A41793	dipeptidyl aminope
14	427	9.1	803	I68600	dipeptidyl aminope
15	427	9.1	865	I54331	dipeptidyl aminope
16	378	8.0	829	T19514	hypothetical prote
17	374	8.0	795	F82588	dipeptidyl-peptida
18	331.5	7.1	759	I38593	fibroblast activat
19	293	6.2	743	T37700	probable dipeptid
20	282	6.0	642	C71137	hypothetical prote
21	269.5	5.7	631	T875007	probable acylamino
22	268	5.7	683	E87495	prolyl oligopeptid
23	255.5	5.4	709	B82580	alanyl dipeptidyl
24	243.5	5.3	657	E70025	probable acylamino
25	236	5.0	622	F71174	hypothetical prote
26	223.5	4.8	632	E75057	peptidase PAB1418
27	217.5	4.6	536	F90299	acylaminoacyl-pept
28	200.5	4.3	591	H72474	probable acylamino
29	185	3.9	569	S74053	probable acylamino

30	183.5	3.9	608	2	F83397	probable peptidase
31	182.5	3.9	674	2	B84381	acylaminoacyl-pept
32	180.5	3.8	721	2	T09631	probable acylamino
33	180	3.8	667	2	A87711	prolyl oligopeptid
34	172	3.7	732	1	JC4655	acylaminoacyl-pept
35	171	3.6	732	1	S07624	acylaminoacyl-pept
36	170.5	3.6	572	2	F72455	probable acylamino
37	163.5	3.5	659	2	F72568	probable acylamino
38	161	3.4	676	2	C97775	acylamino-acid-rel
39	160.5	3.4	732	1	JU0132	acylaminoacyl-pept
40	160.5	3.4	745	2	T33751	hypothetical prote
41	158	3.4	598	2	F84199	prolyl endopeptida
42	157	3.3	614	2	E75094	peptidase [importe
43	155.5	3.3	654	2	AD3183	hypothetical prote
44	155	3.3	629	2	T15945	hypothetical prote
45	154.5	3.3	828	2	G87584	hypothetical prote

ALIGNMENTS

RESULT 1

T32919
hypothetical protein K02F2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32919
R.Maggi, L.; Goela, D.
submitted to the EMBL Data Library, January 1998
A:Description: The sequence of C. elegans cosmid K02F2.
A:Reference number: Z21246
A:Accession: T32919
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-931 <MAG>
A:Cross-references: UNIPROT:O44987; UNIPARC:UPI0000076BD8; EMBL:AF043699; PIDN:AAB97564
A:Experimental source: strain Bristol N2; clone K02F2
C:Genetics:
A:Gene: CESP:K02F2.1
A:Map position: 1
A:Introns: 58/3; 82/2; 131/2; 178/2; 275/3; 322/2; 404/2; 441/1; 464/2; 486/3; 528/2; 5

Query Match	19.6%	Score	922.5	DB	2	Length	931
Best Local Similarity	29.3%	Pred. No.	2.1e-57				
Matches	282	Conservative	145	Mismatches	326	Indels	211
Gaps	41						
QY	33	EPFVYRYSWOLKLLADTRKYHYGMKAPHDHFMVKRNDPDGPHSDRIYVYLA	SCEN	92			
DB	36	EPARFETRSFSQ---	LIDHARSWKTEVRGTTGGFTKISLMRAE---	KDRLNMYAISVVP	89		
QY	93	RENT--LPYSEBP-KTINRAAVLMSWKPLLDLFOATLDYGMV-----	SRE	135			
DB	90	GTNTQSIFSVTIPLELVEKAQVADRK-----	FELKLSGYNDVSIYRMSCRTKTPPSAE	142			
QY	136	EELLRERKRGITGV-GIASYDYHOG-----	SGTFLPQAG-----	167			
DB	143	FTLOCQRORSQVVTGIDYEIRNGKMLMAGDOLFYNPLNEALAAPIAVPDDQS	STEP	202			
QY	168	-----	SGYHVKGDPQG-----	FTQQLRPNLVETSCP-----	NIRMDPKLCP	206	
DB	203	MDISEGTSITGKCSNEAFOSSTVPVTPRIPIKKTSTTEKPTATPTNNFVS	AKVCP	262			
QY	207	ADPQWIAPIHNDIWNISNIVTREERLTVVHNELANMEEDARSAGVATFVLQ	EEFDYSG	266			
DB	263	ADSSLLAVLNKQYVI-----	EKNKLIHRTSSNSKHI	ITN--GVPSVIVQEELERPEG	313		
QY	267	YWCCKAETTPSGKILRIIYENDSEVEI	IHV-----	TSPMLETRRADSFYPTKG	319		
DB	314	IWW-SESKT-----	RLLYEHVNEKVAESQGVNGDPPVAPM-----	KYPRAG	355		
QY	320	TANPKVTFKMEIMDAGRIIDVIDKELIQPEILFEG---	VEYIARAGWTPEGKYAWS	376			
DB	356	TKNAVSTLRM---VILENGKAYDVPDKD-----	EVYIKHCPFYEYITRAGFFSDGTTVMV	407			

QY 377 ILLDRSOTRLQIVLI-----SPELFIPVEDDMERQRLIESVP 414
Db 408 QVMSRDQAQCSLLPIPYDFLLPBELOGSKEONLQLSTDLNNGVWDD--KSHEETMERKP 466
QY 415 DSVT--PLIYBETTDIWINIHDIHFVFPQSHBEE--IBFIFASECKTGFRHLKYITSIL 470
Db 467 RGKLRGTQIHKARNDYWINTHNAIYPLKITDDEHPMYEPIYCLEPNG--SCLALISABE 525
QY 471 KESKYRSGGLPAPSDPKCPKEEIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDS 530
Db 526 DQNGY-----CRTEKLLMAENFSI--NKSMGIWVDEVRELVYVANESH 569
QY 531 PLEHLYVYVNPGEVTRLDRGYSHSCCISOHCDFFIKYSNOK-----NPHCVSLYKL 586
Db 570 FTEWNI--CVSHRTQHQALTESGI-----C--FKSERANGKALDLDHGFACYMT 617
QY 587 S---SPEDDPTCKTEP-W-----ATILDSAGP-LPD--YTPPEIFSFES--TTG 627
Db 618 SVGSPAE---CRPYGFRWKENEVLPSTVAAITVSHGQPDLDHFDSEMFQSKKTG 674
QY 628 FTLYGMLYKPHDLQPGKYPVTLFYGGPQVOLVNNRFGVKYFRLNTLASLGYVWVID 687
Db 675 LMHYAMILRSPNFPYKYVPVHYVYGGGQIWHNDFSMIQYIR---FCRLGYVWVVID 731
QY 688 NRGSHRGKLFEGAFKYNQGTIEDDQVEGLQYLASRY-DFIDLDRVGHGWSYGYLSL 746
Db 732 NRGSAHRTGEFERHIHKMGTVVEVEDQVEGLQMLAERTGGFMDSRVVHGWSYGYMAL 791
QY 747 MALMQRSDIFRVAIAGAPVTLWIFYDTGYTERYMGHPDQNEQGYLGSVAMQAEKPPSEP 806
Db 792 QMLAKHPNIYRAAAGAVSDWRLYDTATYTERYMGYP--LEEHVYGASSITGLVEKLPDEP 850
QY 807 NRLLHLHGFLENVHFAHTSILLFLVRAGKPYDLQIYPOERHSIRVPESGEHYELHLH 866
Db 851 NRLMLVHGLMDENVHFAHTLHVLVDECIKKGKWHELVFPNPERHGVNRNDASIYLDARMY 910
QY 867 YLOE 870
Db 911 FAQQ 914
RESULT 2
A:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-738 <STO>
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:gl3423647; F
A:Gene: CC2154
Query Match 17.9%; Score 842; DB 2; Length 738;
Best Local Similarity 29.9%; Pred. No. 8.1e-52;
Matches 221; Conservative 115; Mismatches 262; Indels 141; Gaps 24;
QY 133 SREELRERKRIGTGTGASDYVHQSGTFLFQAGSGIY--HVKGQPGQFTQQLPRLN 190
Db 94 SEAEKARRERARVARGIVEYSWDRQGRFILVPLDGLYLDADADGKITRLTE----- 146
QY 191 VETSCPNIWMDPKLCPADPDWIAFIHSDNIWISNITVREERLITVHNELANMEEDARSA 250
Db 147 ----TPGDEVDKVPKG--GVSVYVRDQNLIKPVAGGAETALT-----TDGKDLSF 194

QY 251 GVATFVLOEBEDRSGYWKCPKAETTPSGGKILRLIYEENDESSEVEIHHVTSPMLETRA 310
Db 195 GVAEFIGVIRELDRFTFGYWSPDES-----RIVYTRVDESQDIV-----PRA 236
QY 311 D-----SFRPKTKTANPKVTFKMSIMIDAEGRI--IDV--TDKELIOPPELFP 358
Db 237 DIGPGGATVWQRYPRAGRNAVVDLFRDL---ASGKVTAOLDLGANKOI----- 283
QY 359 VEYIARAGMTPEGKYAMSILLDRSOTRLQIVLISPELFIPEVDDVMERQRLIESVDSVT 418
Db 284 --YVARVAMSADGKTVYVQRLSRDQKTLDLLAF-----DAAT 318
QY 419 PL--IYBETTDIWINIHDIHFVFPQSHBEEIIFIFASECKTGFRHLKYITSILKESYK 476
Db 319 GAGKTLITDTPHFIENVNDRPLTDG-----TFLMGSE--KDGNOHLYRYA----- 363
QY 477 RSGGLPAPSDPKCPKEEIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHHL 536
Db 364 -ADGKLI-----QITKGDWPIVGLEG-----VDEARKVAIFASIDTPIERRL 406
QY 537 YVYVYVNPGEVTRLDRGYSHSCCISOHCDFFIKYSNOKNPHCVSLYKLSSP-----ED 591
Db 407 YEVSYAKPKPKALTSAGCWAAKVADNGGAFAGTYSDPKTPSQTALYSADGKRVRWIEE 466
QY 592 DPTCTKTEFWAILDSAGPLPDYTPPEIFSPESSTTGTLYGMLYKPHDLQPGKYPVTLF 651
Db 467 NKLAEGHPY---PYAANLPQ---PEFGSLKAADGETLHVEILKPIGDFPAKKYPAIVS 519
QY 652 IYGGPQVOLVNNRFGVKYFRLNTLASLGYVYVVDNRGSHRGKLFEGAFKYNQGTIEDD 711
Db 520 YTGGAHQARVMQKWHSPSE---RTYLEAGYVIFKLDNRGSGNRSKAFKMRALDRKLTVEV 576
QY 712 DDQVEGLQYLASRYDFIDLDRVGHGWSYGYLSIMALMQRSDIFRVAIAGAPVTLWIFY 771
Db 577 EDQLLGAKELASQ--PYVDADKLGVMGWSYGGFMALMLLTAEPTPFKAGAAGAPTEWSLY 635
QY 772 DTGYTERYMGHPDQNEQGYLGSVAMQAEKPPSEPNNRLLHLHGFLENVHFAHTSILLSF 831
Db 636 DTATYTERYMGKPDENKAGYAYSDINNRIKUL--APGSLLLHLHGMAADNVIFENSTRMLAA 693
QY 832 LVRAKRPYDLQIYPOERHS 850
Db 694 LQRKAILFEMAMYPGERHS 712
RESULT 3
JC5142
X:Pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia
C:Species: Xanthomonas maltophilia
C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5142
R:Kabaehima, T.; Ito, K.; Yoshimoto, T.
J. Biochem. 120, 1111-1117, 1996
A:Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression
A:Reference number: JC5142; MUID:97164011; PMID:9010758
A:Accession: JC5142
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <KAB>
A:Cross-references: UNIPROT:P95782; UNIPARC:UPI00000B5F8A; DDBJ:D83263; NID:gi753196; P
C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl rei
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptidase hydrolase; membrane bound
F:4-18/domain: transmembrane #status predicted
F:610/Active site: Ser #status predicted
F:685/Active site: Asp #status predicted
F:717/Active site: His #status predicted
Query Match 16.0%; Score 754; DB 2; Length 741;
Best Local Similarity 28.4%; Pred. No. 1.5e-45;
Matches 214; Conservative 120; Mismatches 281; Indels 138; Gaps 26;

Qy 133 SREEELLRRKRIGTV-GIASYDHYHGGSTGTFLOAGSG--IYHVKGQGGPGGTFQOPLRN 189
Db 93 SDEKARRERQRIAAWTGIVDQWSPDAQRLLFLPGGELYLDLKQEGKAARQL-----147
Qy 190 LVETSCPNIRMDPKLCPADPDWAFIHSNDIWIINIVTRERRLLTYVHNELANWEEADARS 249
Db 148 ---THGEGFATDKLSPKG-GFVSPIRGNLWIDLASGRQMOLT-----ADGST 193
Qy 250 A---GVATFVLOEEDRYSGYWCPCAEETTPSGGKILRLIYEENDESEV-----EIIHV 300
Db 194 TINGGIAFEVADEMDRHTGYWAPDDSA-----IAYRIDESPVPQKREYEVAD 244
Qy 301 TSPMLRETRADSFYRPTKGTANPKVTFKMSIMIDAEGRIDVIDKELIOPFELPEGVE 360
Db 245 RTDVIQ-----RYPAAGDANQVKLGWISPAEQATQWIDLKQDI-----287
Qy 361 YIARAGWTPEGKVAWSILLDRSOTRLQIVLISPELFIPEVDVWVERORLIESVPDSVTPL 420
Db 288 YLARVNVWRDPQLSFG-RQSRDQKKLDLVEVTL-----ASNQOR-----325
Qy 421 IIVEETTDIWINIHDIHFVFPQSHHEEIEFIPASECKTGPRHLYKITSILKESKYKRSG 480
Db 326 VLAHETSPWVPLHNSLRF-----DDGSILWSSE-RTGFOLHLYRI-----DSGKAA--372
Qy 481 GLPAPDFKCPKEIEAITSGEWEVLGRHGSNIQVDEVRRLVYPEGTKDSPLEHLYVVS 540
Db 373 -----ALTHGNWSV-----DELLAVDEKAGLAYFRAGIESARESQIVAVP 412
Qy 541 YVNPGEVTRLDRGYSHSCCISOHCDFPISKYKQKQNPCHCVSLYKLSPPDDPTCKTKEF 600
Db 413 -LQGGQPORLSKAPGMSHASFARNASVVDWSNNSTPQTELFRANG-----EK 461
Qy 601 WAIL--DSAGPLPDYT-----PPEIFSESTGTFT-LYGMLYKPHDLQPKKYPTVL 650
Db 462 IATLVENDLADPKHPYARYEARQRPVEFTLTAADGKPLNLYVIKIPAGDPDAKRYPAV 521
Qy 651 FIYGGQVQVNNRFGVKYFRLNT--LASLGVVVVVVIDNRGSHRGKLFEGAFKYMGOI 709
Db 522 YVGGPASQTVTDSWFERGDHLPNQYLAQGGYVVFSLDNRGTGPRGRDFGALYKQGTV 581
Qy 710 EIDDOVEGLQYLASRYDFIDLRVGHGWSYGGYLSLMLMQRSDIFRVAIAGAPVTLWI 769
Db 582 EVADQLRGVAWL-KQOPWVDPAIRIGVQWSNGYVMTLLAKASDSYACGVAGAPVTDWG 640
Qy 770 FYDTGTYRMGHDPDQNEOGYVLGSVAMQAEKFPSEPNRLLLHGFLENVHFAHTSILL 829
Db 641 LYDSHYTERTYMDLPAENDAGYREARVLTHIEGLRSP---LLLIHGMADDNVLTNLSLM 697
Qy 830 SFIVRACKPYDLQIYQPERHSIRVPESGHEVEL 862
Db 698 SALQKRGQPELMTYPGAHKHGLSGADALHRYRV 730

RESULT 4

S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C:Species: Flavobacterium meningosepticum
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S66261
R:Kobashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A:Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from *P. fluorescens* Pf-5
A:Reference number: S66261; MUID:95314307; PMID:7793970
A:Accession: S66261
A:Status: preliminary
A:Molecule type: DNA
A:Molecule type: mRNA
A:Residues: 1-711 <KAB>
A:Cross-references: UNIPROT:Q47900; UNIPARC:UPI00000B2115; EMBL:D42121; NID:G577283; PDB:1YVH
A:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptidase hydrolase

Query Match 13.28; Score 620; DB 2; Length 711;
Best Local Similarity 25.28; Pred. No. 5.2e-36;

Matches 195; Conservative 134; Mismatches 284; Indels 160; Gaps 26;
Qy 145 IGTGVIASDYDHYG-----SGTF-----LFOAGS-----GIYHVWDG 176
Db 54 IEPTGIKYSYKTSQKKNIVDGSFGQYTFNSDESKILLQKSSQSIYRHSFLGKFEVKD- 112
Qy 177 GPOGFTQOPLRP-NLVETSCPNIRMDPKLCPADPDWAFIHSNDIWIINIVTRERRLLTY 235
Db 113 -----LKSTVTVSLNANWIOEPKPSF-DGSKVAFIADNNLFYODLNTGKITQLTT 162
Qy 236 --VHNELANWEEADARSAGVATFVLOEEDRYSGYWCPCAEETTPSGGKILRLIYEENDES 293
Db 163 DGKKNELIN-----GLGDWVYEEFGHADIYQW-----NKAGDALVFREDKVP 208
Qy 294 EVEL-IHVTS--PMLRETRADSFYRPTKGTANPKVTFKMSIMIDAEGRIDVIDKELIQ 350
Db 209 EINIPIYQNLKPLMT-----YKYPKAGEENSAVTAYLYQL-----SSGK-----249
Qy 351 PFEILPEGVE--YIARAGWTPEGKVAWSILLDRSOTRLQIVLISPELFIPEVDVWVERQR 408
Db 250 SAQLNFGSSSEKYIYQIFQTNANDEIVVATANRHNQKVDLLKVNTK-----295
Qy 409 LIESVPDSVTPLLIYBEETTDIWINIHDIHFVFPQSHHEEIEFIPASECKTGPRHLYKITS 468
Db 296 -----TAAVSKLFTETONAMETDNLWTF-----LDNSFLWASE-RDGRHLYWYDA 343
Qy 469 ILKESKYKRSGGLPAPSDFKCPKEIEAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTK 528
Db 344 AGKLKK-----QVSKGDWEIINYGNPKTKE-----VVIQTTE 377
Qy 529 DSPLEHLYVVS--YVNPGEVTRLDRGYSHSCCISOHCDFPISKYKQKQNPCHCVSLYKL 586
Db 378 KGSINK---VVSKLINTGKTQLLSNAEGNSAASFSTFNFIINTSSTAKVPTKYLKDA 434
Qy 587 SSPE-----DDPTCKTKEFWATILDSAGPLPDYTPPEIFSESTGTGTYLGMLYKPHD 639
Db 435 NGKDVKELQNDDLLNKLKS-----DNFAKEFITIPNAAGDQWAMWIKPN 482
Qy 640 LQPKKYPTVLFIYGGPQV 699
Db 483 FDPAKKYVPMFQYSGFGSQVANSWDGNGWFDMLAQKGLVVCVDRGTGFRGTGYK 542
Qy 700 GAFKYMGOIIDDQVEGLQYLASRYDFIDLRVGHGWSYGGYLSLMLMQRSDIFRVA 759
Db 543 KVTYNLKGVEIEDQITAAKWLGNQ--SYVDKSRIGIFGWSYGGYLSLMLMQRSDIFRVA 601
Qy 760 IAGAPVTLWIFDTGTYRMGHDPDQNEOGYVLGSVAMQAEKFPSEPNRLLLHGFLEN 819
Db 602 IAVAPVNWRFYDSIYTERFLQTPQENKQGYDLNPTTYAKLL---KGKFLIHGTADDN 658
Qy 820 VHPAHTSILLSFLVRACKPYDLQIYQPERHSIRVPESGHEVELHLLHYLQENL 872
Db 659 VHFQNSMEFSEALIQNKQDFMAYPDKNHSIIGNTRPQLYEKMTNILENL 711

RESULT 5

CDHU26

dipeptidyl-peptidase IV (EC 3.4.14.5) - human

N:Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004

C:Accession: S24313; B42408; B61136; S59510; I56154; S59857; S15520

R:Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehara, Y.

Biochim. Biophys. Acta 1131, 333-336, 1992

A:Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a

A:Reference number: S24313; MUID:92329551; PMID:1352704

A:Accession: S24313

A:Molecule type: mRNA

A:Residues: 1-6, '1', 8-766 <MIS>

A:Cross-references: UNIPROT:P27487; UNIPARC:UPI000016AE80; EMBL:X6708; NID:G35335; PDB:1YVH

R:Darmoul, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A.

J. Biol. Chem. 267, 4824-4833, 1992

A:Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer

Db 277 IKYKPGNPNPSVSLFVADLNSSNFSLMNEPLAEP-----VQNVLMVNTS----- 326
QY 373 YAWSILL---DRSOTRLQIVLISPEL--FIPVEDDVMER-----ORLIESVPDSTVPLII 422
Db 327 ---SVLVQFTNRNSTCITARLLDTELKSIHTVKTCEBEGWEVQOSAKMFP--LNNSLV 381
QY 423 YEETDWINIHDIHFVFPQSHBEEIEFTASECKTGPRHLYKITSILKESKYRRSGGL 482
Db 382 WENWSD---GYFDILALDDYNHLAIFP-----NGSS 410
QY 483 PAPDFKCPKEETAITSGEWEVLGRHGSNIQVDEVRRLVPEGTKDSPLHLYVVSIV 542
Db 411 P-----IYLTSGAWDT---DGPIDHGDGFGNVYPLATLKDSTERHLYYS-L 454
QY 543 NPGEVTRLTRDGSY---HSCCISQHCDFFTSKYSNOKNPHCVSLYKLSPPDDPTCKTKE 599
Db 455 DTLEYGITDNGEDEGYSTSFSPFGDFYVLYNVHGDVP---WQELASTKDQCYLSLE 510
QY 600 FWATILDSAGPLPYTPPEI-----PFSBSTTGTLYGMLYKPHDLQPKKYPTVLFYGG 655
Db 511 TNSRLKQO---LSSITLPSVEYKGLTFNDTT-FNF--MERRPRNFDVNNKYPVLFYAYGG 564
QY 656 PQVQLVNNRFKGVKYFRLNTLASLG-----YVVVIDNRGSGHRLKFEKAFYKMGQI 709
Db 565 PGOQOV-----AKLFRVDYQAYLASHPDPEFIVTLVLDGRGTGFGNGAFRYSVSRHLGEW 618
QY 710 EIDQVEGLQYASRGYDFIDLDRVGIHGSYGGYLSMALMORSDFRVAIAGAPVTLWI 769
Db 619 ESYDQOQAGKFWAD-LFVDENHVGIVGWSYGGYLTKLTL-ETQDVFSGYNAVAPVDWR 676
QY 770 FYDTGYTERYMGHPDQEOGYLGSVAMQAEKPPSEPNRLALLHLLHLYQENLGSRIAAL 879
Db 677 LYDSVYTERYMDLPQYNKEG-YKNSQIHDIYERF-KQLKRFVFAHGTGDNDVHFQHSNHL 734
QY 830 SFLVRAG-KPYDLOIYQERHSIRVPESGEHLYLHLLHLYQENLGSRIAAL 879
Db 735 DGLNLANCYNVMAVFPDPSAHSI-----SYHNASLSIYHRLSEWIGDALGRI 781

RESULT 7

S23752
diptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
N:Alternate names: CD26 alpha subunit; THAM alpha subunit
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S23752; A46465; A56030
R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
J. Biol. Chem. 267, 2200-2208, 1992
A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di
A:Reference number: S23752; MUID:92129288; PMID:1370813
A:Accession: S23752
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-760 <MAR>
A:Cross-References: UNIPARC:UPI0000172A2D; EMBL:X58384
R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
J. Immunol. 147, 447-454, 1991
A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (diptidyl peptidase
A:Reference number: A46465; MUID:91302787; PMID:1712807
A:Accession: A46465
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <VIV>
A:Cross-References: UNIPARC:UPI0000172A2E
A:Experimental source: M14.T thymoma cells, Swiss nu/nu
A>Note: sequence extracted from NCBI backbone (NCBIP:42236)
R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
Biochemistry 33, 15204-15214, 1994
A:Title: Structure of the mouse diptidyl peptidase IV (CD26) gene.
A:Reference number: A56030; MUID:95092780; PMID:7999781
A:Accession: A56030
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 746-760 <BER>
A:Cross-References: UNIPARC:UPI0000172A2F; GB:U12620
C:Genetics:
A:Gene: CD26
C:Superfamily: diptidyl-peptidase IV
C:Keywords: diptidyl-peptide hydrolase; glycoprotein; transmembrane protein
F:213,223,315,514,679/Binding site: carboxylate (Asn) (covalent) #status predicted
F:624,702,734/Active site: Ser, Asp, His #status predicted

Query Match 11.0%; Score 518.5; DB 1; Length 760;
Best Local Similarity 25.0%; Pred. No. 1e-28;
Matches 203; Conservative 107; Mismatches 306; Indels 197; Gaps 37;

QY 131 MYREELLRRKRKIGTVGI-----ASYVHQSGTGLP-----QAGS 168
Db 67 LYKQENILLNLAHEGNSIFLENSTFESFGYHSVSPDLFLVLLYNYVYKWRHSYTSY 126
QY 169 GYHVKDGQGGTQOQPLRNLVETSCPNIRMDPKLCPADPDWI-----AFIHSND 219
Db 127 NIYDNK-----ROLITEEKIPN-----NTQWITWSPEGHKLATVWKN 165
QY 220 IWNISNIVTREERRLTYVHNELANMEEDARSAGVATFVLOEE-FDRSYGVWCPK----- 272
Db 166 IYKVEPHLPSHRI-----STGEENVYNGITDWDVVEEVFGAYSALWSPNNTPLAY 219
QY 273 AETPESGKILRLIYEENDESEVEIITHVTSPLMELTRADSFRYPTGTANPKVTPKMEI 332
Db 220 AQFNDTGVPLEIYSF-YSDS-----LQYPTWIPYKAGAVNPTVKPFI--V 265
QY 333 MIDAEGLIDVIDKELIQPEILFEGVEYIARAGTTPCKYA-----WSILLDRSQ 383
Db 266 NIDSLSSSSAAPIQIPAPASVA-RGDHYLDCVWATEERISLQWLRRIQNYVNAICDY 324
QY 384 TRLOILISPELFTPVEDDVMERQRLIESVPDSVTLPIIYEETDIWINIHDIHFVFPQS 443
Db 325 DKINLTWNCPS-----EQOHV-----ENSTTGWG--RFRPAEHP 358
QY 444 HEERIFIFASECKTGPRHLYKITSILKESKYRRSGGLPAPSDPK-CPKEEIAITSGE 502
Db 359 TSDGSSFYKIISDKGKYKICHF-----PKDKKDCITF-----ITKGA 395
QY 503 WEVLGRHGSNIQVDEVRRLVPEGT-KDSPLEHLYVVSYPVNGEVTRLTRDGYSHSCCI 561
Db 396 WEVI-----SIEALTSYLYISNQYKEMPGGGRNLYKI-----QLTD--HTNVKCL 439
QY 562 S-----OHCDFPISKYSNOKNPH---C-----VSLYKLSPPDDPTCKTKEFWATILDSAG 609
Db 440 SCDLNPCKQYVAVSFSEKAKYQLGCGWGPGLPLYTLHRSTHKLRLVLE-----DNSA 493
QY 610 ---PLPDYTPPE---IFSFSSTTGTLYGMLYKPHDLQPKKYPTVLFYGGPQVQLVNN 663
Db 494 LDRMLQDVQMPKSKLDLFLVNETRF-WYQMLPPH-FDKSKYPLLLDLYAGPSCS---- 547
QY 664 RPKGVKVPRLN---TLASLGYVVV-IDNRGSGHRLKFEKAFYKMGQIEIDDOVEGLQ 719
Db 548 ---KADASFLNWTATYLASTENIIVASFDGRGSGYQGDKIMHAINRRLGLEVEDEIAR 605
QY 720 VLASRYDFIDLDRVGIHGSYGGYLSMALMORSDFRVAIAGAPVTLWIFDYDTGYTER 779
Db 606 QFV-KMGFVDSKRVAVIWNYSYGGYVTSWLGSSGVFKGIAVAPVSRWEYDYSYTER 664
QY 780 MG---HPDQEOGYLGSVAMQAEKPPSEPNRLALLHLLHLYQENLHLYQENLHLYQEN 837
Db 665 MGLPIPEDNLDRYNSVMSRAEHF--KQVEYLLIHGTADDNVHVFQSSAQISKVLVDAGV 722
QY 838 PYDLOIYQERHSIRVPESGEHLYLHLLHLYQENLHLYQENLHLYQENLHLYQEN 870
Db 723 DFOAMWYTDDEHGIASSTAHCHYSHMSHFLOQ 755

RESULT 8

A39914
diptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat

N;Alternate names: GP110; membrane glycoprotein 110K; OX-61
N;Contains: dipeptidyl-peptidase IV, soluble form
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C;Accession: A39914; A33315; A60730; A42203; S38949; A31781
R;Hong, W.; Doyle, D.
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
A;Title: cDNA cloning for a bile canalicular domain-specific membrane glycoprotein of rat
A;Reference number: A39914; MUID:88068516; PMID:3479775
A;Accession: A39914
A;Molecule type: mRNA
A;Residues: 1-792 <HO>
A;Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:
R;Ogata, S.; Misumi, Y.; Ikehara, Y.
J. Biol. Chem. 264, 3596-3601, 1989
A;Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA an
A;Reference number: A33315; MUID:89123496; PMID:2563382
A;Accession: A33315
A;Molecule type: mRNA
A;Residues: 1-37, 'A', 39-182, 'I', 184-331, 'T', 333-351, 'C', 353-393, 'V', 395-561, 'L', 563-623,
A;Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:G203973; PIDN:AAA41096.1; PID:
A;Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
A;Accession: B33315
A;Molecule type: protein
A;Residues: 1-20,35-54,427-443;505-509;511-520;530-538;593-600;602-608;618-627 <HO>
A;Cross-references: UNIPARC:UPI0000172A2E; UNIPARC:UPI0000172A30; UNIPARC:UPI0000172A31;
A36; UNIPARC:UPI0000172A37
R;McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.
Hepatology 11, 534-544, 1990
A;Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop
A;Reference number: A60730; MUID:90228896; PMID:1970322
A;Accession: A60730
A;Molecule type: protein
A;Residues: 28-47, 'XX', 50-53,55-58 <MC>
A;Cross-references: UNIPARC:UPI0000172A38
R;Ogata, S.; Misumi, Y.; Teuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
Biochemistry 31, 2582-2587, 1992
A;Title: Identification of the active site residues in dipeptidyl peptidase IV by affini
A;Reference number: A42203; MUID:92190188; PMID:1347701
A;Accession: A42203
A;Molecule type: protein
A;Residues: R', 625-630, 'X', 632-648 <OC>
A;Cross-references: UNIPARC:UPI0000172A39
R;Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
A;Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl p
A;Reference number: S38949; MUID:94128239; PMID:7903271
A;Accession: S38949
A;Status: preliminary
A;Molecule type: protein
A;Residues: 281-302 <IWA>
A;Cross-references: UNIPARC:UPI0000172A3A
R;Hong, W.; Doyle, D.
J. Biol. Chem. 263, 16892-16898, 1988
A;Title: Membrane orientation of rat gp110 as studied by in vitro translation.
A;Reference number: A31781; MUID:89034185; PMID:3182821
A;Accession: A31781
A;Molecule type: mRNA
A;Residues: 1-40 <HO>
A;Cross-references: UNIPARC:UPI0000172A3B
C;Comment: This protein is localized to the bile canalicular, which is the apical domain
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidyl-peptidase, glycoprotein; homodimer; liver; serine proteina
F;1-28/Domain: signal sequence #link MATS #status experimental <SIG>
F;1-28/Domain: intracellular #status predicted <INT>
F;1-28/Domain: transmembrane #status predicted <TN>
F;29-792/Domain: extracellular #status predicted <EXT>
F;29-792/Domain: propeptide #link MATS #status experimental <PRO>
F;35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <MATS>
F;83,90,148,217,227,319,421,686/Binding site: carbohydrate (Asn) (covalent) #status pred
F;631/Active site: Ser #status experimental

F;709,741/Active site: Asp, His #status predicted
Query Match 10.9%; Score 510.5; DB 1; Length 792;
Best Local Similarity 25.7%; Pred. No. 4e-28;
Matches 182; Conservative 105; Mismatches 287; Indels 133; Gaps 30;
QY 206 PADPWIAFIHSD-----IWSNIVTEERRLTYVNEELANMEEDARSAGVATFVLQEB- 260
DB 147 PNTQITWISQREGHKLAVYKNDIYVLEPHLP-SHRTTSTGKENVFNGINDWVVEEI 205
QY 261 FDRSYGWCPCAEETTPSGGKILRLIYEENDESEVEIHVTSPLMTRADSFRYPKCTG 320
DB 206 FCAYSALMW-----SPNGTFLAVAQNDGTGVLIEYSFYSDLSQTPKTVWIPYKAGA 259
QY 321 ANPKVTFKMSIMDAEGRIIDVIDKBLIOFFELFEGVEYIARAAGWTPEGKYAMSLILD 380
DB 260 VNPVTVKFET--VNTDSLSTTTTTPMOTAPASVT-TGDHYLCDVAWVSEDRIS----- 310
QY 381 RSQTRLQIVLISPFLFIPVEDDVNVERORLIESVPDSVTPLIIYEETTDIWINIHDFHV- 439
DB 311 -----LQWLRRIQNY--SYMAICDYDKTNLVMNCFTTQEHLE 345
QY 440 -----FPQSHHEBIEFIFASECKTGRHLYKITSLKESKYKRSSGGGLPAPS 486
DB 346 TSATGWGVRFRPAEPHFTSDGSSFFKIVSDKDYKH---ICQFQDKRK----- 390
QY 487 DFKCPKEBIATTSGEVWLGHRGNSNIQVDEVRRLVYFEGT-KDSPLEHHLHYVSVNPG 545
DB 391 -----PQDCTFITKAMEVI-----SIEALTSYLYISNEYKEMPGRNLYKI----- 435
QY 546 EVTRLTDRGYSHSCCIS-----OHCDFP---ISKYSNOKNPHC-----VSLYKLSPPEDDP 593
DB 436 ---QLID--HTNKKLSCDLNPERCOYYSVLSKEAKYVOLGCRGCPGLPYTLTHRSDDQK 490
QY 594 TCKTKFWATILDSAGPIPDYTPPE---IPSFESTTGTLTYGLMYLKPDLQPGKYPTVL 650
DB 491 ELRVLED-NSALDKM--LDVQMPKSKLDIVLNETRF-WYQMLLPH-PDKSKYPLLI 545
QY 651 FYVGGQVQLVNNRFGVKYFRLN---TLASLGYVVVV-INNRSGHRGLKFEAGFYKM 706
DB 546 DVYAGPCSQ-----KADAAFRFNWATYLASTENIIVASFDRSGSGYQGDKIWHAINKRL 599
QY 707 GQIEIDDOVEGL-QYLASRYDFILDRVGIHWSYGGYLSLMLMQRSDIFRVAIAGPV 765
DB 600 GTLLEVEDQIEAARQFL--RMGFVDSKQVAIWGVSYGGYVTSVLGSGSGVFKCGIAVAPV 657
QY 766 TLMIFYDTGTYERYMG--HPDQNEQGYLGSVAMQAEKPPSEPNRLLLLHGLFDENVHFA 823
DB 658 SRWEYVDSVYTERYMGLTPTEDNLDHYNSTVMSAENF--KQVEYLLIHGTADDNVHFQ 715
QY 824 HTSILLSFLVRACKPYDLQIYQERHSIRVPESGEHVELHLLHLYQE 870
DB 716 QSAQISKALVDAGVDFQAMWYTDHGIASSTAHOHIYSHMSHFLOQ 762

RESULT 9

A30107
dipeptidyl aminopeptidase B (BC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YHR028c
C;Species: Saccharomyces cerevisiae
C;Date: 07-Jun-1990 #sequence revision 30-May-1997 #text change 09-Jul-2004
C;Accession: S46780; A30107
R;Du, Z.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of S. cerevisiae cosmid 8082..
A;Reference number: S46773
A;Accession: S46780
A;Molecule type: DNA
A;Residues: 1-818 <DUZ>
A;Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031ASF; EMBL:U10399; NID:G500689; PI:
R;Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an i

A:Reference number: A30107; MUID:89174971; PMID:2647766
A:Accession: A30107
A:Molecule type: DNA
A:Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRRLET', 189-199, 'N', 201-365, 'DFKGRKERK', 376-574
A:Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A:Note: the authors translated the codon ACC for residue 572 as Asn
C:Genetics:
A:Gene: SGD:DAP2; STB13; MIPS:YHR028C
A:Cross-references: SGD:S0001070; MIPS:YHR028C
A:Map position: 8R
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacuole
F:30-45/Domain: transmembrane #status predicted <TM>
F:63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.3%; Score 484.5; DB 1; Length 818;
Best Local Similarity 26.08; Pred. No. 3e-26;
Matches 183; Conservative 95; Mismatches 250; Indels 177; Gaps 35;

QY 212 IAFTHSNDIWSNIVTREERLTYVHNE-----LANMEEDARSAGVATFVLOEE-FDRYSG 266
DB 201 IAYQDNIIYISAISKKTIRA--VTNDGSSFLNGKPD-----WVYEEVFEDDKA 250
QY 267 YWCPKAEPTPSGKILRLIYENDESEV-EIHVTSPLMLETTRAD-----SPRYPTKG 319
DB 251 AAWGPTGD-----YLAFLKIDSEVGEFT---IPYVQDEKDIYPMRSIKYPRSG 298
QY 320 TANPKVTFKSEIMIDAGRIIDVIDKELIQP--FEILFEGVEYIARAGWTPEKQAWSI 377
DB 299 TPNP-----HAELWVYSKMDGTGFHPRISGNKKDGLLITVTVWGNG-----NV 343
QY 378 LL-----DRSOTRLQVLISPELFIPEVDVMEORLIESVPDSVTPLLIYEETDIWINI- 433
DB 344 LVKTTDRSSDILTVFLI-----DTAKTSNVVRN-----ESNGGWWEIT 383
QY 434 HDIFHV-----FPQSHEEIEFIPASECKTGRHLKYKTSILKESKYRSGGLPAPSDF 488
DB 384 HNTLFI PANETFDPRHNGYVDILPIG---GYNHL-----AYPENSNS-----SHY 425
QY 489 KPCKIEIAITSGEWEVLGRHSGNIQVDEVRRLVPEGTQDSPLEHLVYVSYVNPGEVT 548
DB 426 K-----TLTEGKWEV---NGPLAFDSMENRLYFISTRKSSSTERRHYVIDLRSNPEII 475
QY 549 RLTRDGRYS--HSCCISQHCDFISKYSNKNPH-----C-----VSUYKL 586
DB 476 EVTITSEDGVYDVSFSGRRFGLLTYGKVPYQKIVDFHSRKAECCKGNVLGSLYHL 535
QY 587 SSPEDDPTCKTFWATILDSAGPLPDYTPPEIFSFESTTGTLYG--MLYKPHDLQP-- 642
DB 536 EKNE-----VLTKI--LEDYAVPRKSFRELNLGKDEFGKDLVNSVEILLPND 580
QY 643 -----GKKYPTVLFIYGGPOVLVNNRFGVKYFRLN-----TLASLGYYVVVIDNRSGCH 693
DB 581 FDETLSDHYVPFFPFGNPSQQV-----VKTFSGFNEVVASQNALIVVVDVGRGTGF 634
QY 694 RGLKFEAGFYKMGQIIDDQVEGLQVLASRYDFIDLDRVGHGWSYGYLSLMAQMRS 753
DB 635 KGQDFRSILVRDLGDYEARQDISAASLYGS-LTFVDPQKISLFGMSYGYLFLKLEKDG 693
QY 754 D-IFRVAIAGAPVTLWIFDYTYGTYRMYGHPDQNEQGYLGSV-----AMQAKFPFSEPNR 808
DB 694 GRHFYKYGNSVAPVDMFVDSYVTERMYHTPQENFDGYSVSSHVNTAL-----AQANR 747
QY 809 LLLHGLFDENVHFHFTSILLSLFLVRAG-KPYDLQIYPOERHSIR 852
DB 748 FLAMHGTGDDNVHFQNSLKFLDLDLANGVENYDVHVFDPDSHSIR 792

RESULT 10
A49737
dipeptidyl aminopeptidase (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O5045; protein YOR219c; protein YOR50-9
C:Species: Saccharomyces cerevisiae

C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
A:Accession: A49737; S45451; S60946; S67112; S71721
R:Santa Anna-A, S.; Herskowitz, I.
submitted to the Protein Sequence Database, July 1993
A:Reference number: A49737
A:Accession: A49737
A:Molecule type: DNA
A:Residues: 1-931 <SAN>
A:Cross-references: UNIPROT:P33894; UNIPARC:UPI0000136060; GB:L21944; NID:g347196; PIDN:PIDN
R:Anna-Arriola, S.S.; Herskowitz, I.
Yeast 10, 801-810, 1994
A:Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase
A:Reference number: S45451; MUID:95066382; PMID:7975897
A:Accession: S45451
A:Molecule type: DNA
A:Residues: 1-931 <ANN>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA35119.1; P
R:Gallison, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A:Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
A:Reference number: S60938
A:Accession: S60946
A:Molecule type: DNA
A:Residues: 1-931 <GAL>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1;
R:Boyer, J.; Fairhead, C.; Gallison, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon,
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67104
A:Accession: S67112
A:Molecule type: DNA
A:Residues: 1-931 <BOY>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:Z75127; NID:g1420507; PIDN:CAA99437.1;
A:Experimental source: strain S288C
R:Gallison, F.; Dujon, B.
Yeast 12, 877-885, 1996
A:Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV
A:Reference number: S71713; MUID:96437977; PMID:8840505
A:Accession: S71721
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-931 <GAW>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1;
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C:Genetics:
A:Gene: SGD:STE13; YC11
A:Cross-references: SGD:S0005745; MIPS:YOR219c
A:Map position: 15R
C:Function:
A:Description: involved in processing of alpha-factor prepropheromone
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F:113-150/Domain: transmembrane #status predicted <TM>
F:377/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:785,863,896/Active site: Ser, Asp, His #status predicted

Query Match 9.9%; Score 465; DB 2; Length 931;
Best Local Similarity 21.5%; Pred. No. 9e-25;
Matches 218; Conservative 150; Mismatches 380; Indels 268; Gaps 41;

QY 6 ETEQLGVEIFTACENIESQBPKEPPEYVRYSWSQLKLLADTRKYHGYMAKAPH 65
DB 43 QSNENGHTINIRTEATIDVTDPQ--TPFLQEQYS-----MRPRE 83
QY 66 DMFVKRNDPGPHSDRIYVYLAMS-----GENRENTLFYSBPKTINRAAVLMLS 115
DB 84 SFQF---NDIENQHTHSFPVKNFRWGEWSLPEKRSVLVFTLIALSVLLVILIP 140
QY 116 WKPL-----LDLFOATL-----DYGM 131
DB 141 SKLLPTKITRPTKSAGDSSLGKRSFSIENVLNGDFAIPEDTFHFIDPQRLLGQSDPGL 200
QY 132 ISREELLREKRIG-TVGASDYDYHCGSGTFLFQ----- 165
DB 201 YPTTEIDGHTNFIKQLFDETFEVNIGNRFYEGVEFTVTVQINLYKDKLIFGTNLE 260

C;Superfamily: dipeptidyl-peptidase IV	
Query Match	9.6%; Score 451.5; DB 2; Length 799;
Best Local Similarity	23.0%; Pred. No. 6.5e-24;
Matches 195; Conservative 129; Mismatches 298; Indels 227; Gaps 38;	
QY	78 PHSDRIYVLAMS-----GENRENTLFYSEIPKTIINRAAVLMLSKWPKLL-----DLF 123
Db	116 PSADRKYFAMMDHAFNPMGNFQNETFHLKIVNNNERL-----NPLPFEVEELPRELS 168
QY	124 QATLDYGMYSREELLREKRIGTVGIASVDYHOGSGTFLFOAGSGYHYHVKDGGPOGFTQ 183
Db	169 DSRITVDIGLKEESV-----IQAPKWKGNFDFVFNESNKIYY-----207
QY	184 QLRPNLVETSPNIRMDPKLCPADPOMIAFIHSNDIWIISNIVTREERRITVYHNELANN 243
Db	208 -----QSSP-----ELEGTRVNSGHEHT 226
QY	244 EEDARSAGVATFVLOEE--FDRYSGYWMCPKAETTPSGGKILRLIYEENDESEVEIIHVTS 302
Db	227 VD-----GLFDWIYEEEIFGRKAMWSTK-----GDQAYASYDNHLTKVSL--KTY 273
QY	303 PMLETRAD--SPRYPKT-----GTANPKVTFKMSIEMIDAEGRITDVIDIKELIQPEIL- 355
Db	274 HRLPEPIDTNRHYKTPAKVLPYTLISWNKTE-----QSRQLDVQLKDSLSYHYLLA 328
QY	356 -----PEGVEYIARAGWTPEGKYAWSILL-----DRSOTRLQIVLISPELFIPEVDVME 405
Db	329 VKWLEINGTEQLVSV-WT--NRYQNEVALTICDWDTAICRLEPE-----369
QY	406 RQRLESVPDSTVPLIIEETDDIWINIHDIHFHVPQSHHEEIEFIFASECKTGFPHLYK 465
Db	370 -----YKASKRWVT--HDDFHSI--TSFEDTLFLLP-----HDKR 402
QY	466 ITSILKESYKSSGGLPAPSPDFKCPIKEEIAITSGEWELGRHGSNIQVDEVRRLVYFE 525
Db	403 DNAFOQVASLRLSHGQLRTPK-----FLNLGEYDVTSSINGINKET-----RTIPFH 448
QY	526 GTKDSPLEHLYVVSVPNPEVTRLTDRGYSHSCIS---QHCDFFISKYSNQNPKHCVS 582
Db	449 AAAPKPSHRSLSFYS-----LADESRNSAYCISCSIKNCTWAOQDDQMKTAIVS 499
QY	583 LYKLSGSPED-----DPTCKTK-----EFWATILDSAGPLPDYTPPEIFSFESTTG 627
Db	500 CKGPAAPHATAIVNLTRMDSDKKTEHANLLYDKTYQNRVEEAG--LPVIAKETI---KISDD 555
QY	628 FTLYGMLYKPHDLQPKKY---PTVLFYGGPQVQLVNNRFGKGVKFRNLNTLASLGYVVV 684
Db	556 FDALIKLSIPKDIYNRDKHQAIPLIHVHVGSDNQ---NTKEATQIGIEEVVASASQAAI 612
QY	685 V-IDNRGSGCHRGKFEKFAFKYKMGQIEIDDOVEGLQYLASRY-DFIDLDRVGIHGWSYGG 742
Db	613 LRIDGRSGGRGWKYSALYQGLGIVEVEDQIKAIKVVLLRYRLHLLDARRVAVFGWSYGG 672
QY	743 YLSLMALMQRSD--IPRAVATAGAPVTLWIFDGTGYTYRMYGHPDQNEQGYLGSVAMQAEK 801
Db	673 FMTLSMNEAPEOFFKCAVSAPVTFNPAIDATYTYRYWG--DAPLESY--SDVTKKLDN 728
QY	802 FPSPENRLLHGFLENVHFPAHTSILLSFLVRAGKPDQLQYPOERHSIRVPSGGEHYE 861
Db	729 FKS--TRLLMHGLDDNVHFNQNSALLDELQNRGVDFDLMVIFNQAHSLSRYS--HVV 784
QY	862 LHLHLYLOE 870
Db	785 GKMTFLRQ 793
RESULT 12	
T25173	
hypothetical protein T23F1.7a - Caenorhabditis elegans	
C;Species: Caenorhabditis elegans	
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004	
C;Accession: T25173	

QY	166 -----AGSGIYHVKDGPGQFTQOPLRPNLVETSPNIRMDPK-----LCRAD 208
Db	261 SEFRHSSKGYWIKD-----LNTGNIEPILPPEKSDDDNYELGLSKLSYAHFSPA- 309
QY	209 PDWIAFIHENDIWIISNIVTREERRITY--VHNELANMEEDARSAGVATFVLOEE--FDRYSG 266
Db	310 YNYIYFVYENNLFLQOVNSGVAKKVTEDGSKDIENAKPD-----WIYEEVLASDQA 361
QY	267 YMWCPKAETTPSGGKILRLIYEENDESEVEIIHVTSPLMTRADS--FRYPKTGTANPKV 325
Db	362 IWW-----APDASKAVAFENDISVDDIRLNRYTN--MNEAYLSDKIKYKPGFQNPQF 414
QY	326 TFKMSIEMIDAEGRITDVIDIKELQPEILFEGVEYIARAGWTPEGKYAWSILLDRSOTR 385
Db	415 DL-----FLVNLQNGIYISNTG--GOKOSILYNG-----KWISPTFRPEI--TDNSKI 461
QY	386 LQIVLSPBELFIPVEDDVMERQRLIESVPDSTVPLIIEETDDIWINIHDIHFHVPQSH 444
Db	462 LDVKVY-----DIPSSQML--TVRNTSNLNF-----NGWIEKTKDILSIPKPE 503
QY	445 EEBIE--FIPASECKTGFPHLYKITSILKESYKRSGLPAPSPDFKCPIKEEIAITSGE 502
Db	504 LKRMGYGYIDIHADSRGFSHLFYPTVF-----AKEPIQITKGN 542
QY	503 WEVLGRH--GSIQVDEVRRLVYFEGTKDSPLEHLYVVSV-----NPGVET 548
Db	543 WEVTNGIVGYEYETD---TIFFTANEIGVMSQHLYSISLTDSTTQNTQTSQSNP---594
QY	549 RLTRDGYVSHSCISQHCDFIFISKYSNQKNP-----HCVSLYKLSGSPEDD 592
Db	595 --SDKYDFYDFELSSSARVAISKGLPDTPIKAVGPLTRVLNVAEIHDDSIQLTKDE-- 650
QY	593 PTCCKTEFWATILDSAGLPDYTPPEIFSFESTT---GFTLYGMLYKPHDLQPKKYPTV 649
Db	651 ---KFKX-----KIKNYDLP--ITSYKTMVLDDGVEINYIEIKPANLNPKKYPII 696
QY	650 LFIYGGPQVQLVNNRFGKGVKFRNLNTLASLGVVVVVVINDRSGCHRGKFEKFAFKYKMGQI 709
Db	697 VNIYGGSGQTFT--KSLAFEQAVVSGLDVIVLQIEPRGTGKGKGSFRSWAREKLYW 754
QY	710 EIDQVE--GLOYLASRYDIDLDVRGIHGWVGGVLSMAL--MORSDIERVAATAGAPVTL 767
Db	755 EPRDITEVTKKTFORNSQHIDSKAIWGSYGGFTSLKTYVELDNGDTFKYMAVAFVTN 814
QY	768 WIFYDTGYTRYMGHPDQNEQGYIYLSGVAMQAEKPPSEPNRLLHLLHGFLENVHFPAHTSI 827
Db	815 WTLYDSVYTRYMWNQSENHEGYFEVSTIQNFKFSFES--LKKLFVIFHGTFDDNWHIQNTR 873
QY	828 LLSFLVRAG--KPYDLQIYPOERHSIRKVPESGHEHLLHLYLQENLGSRIAALKVI 882
Db	874 LVDQNLGLTNYDMHIFPDSDHSIRYHNAQRIVFQKLYYWLRFDAFERFDNTEVL 929
RESULT 11	
T25174	
hypothetical protein T23F1.7b - Caenorhabditis elegans	
C;Species: Caenorhabditis elegans	
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004	
C;Accession: T25174	
B;Wilkinson, J.	
submitted to the EMBL Data Library, October 1996	
A;Reference number: 219990	
A;Accession: T25174	
A;Status: preliminary; translated from GB/EMBL/DBJ	
A;Molecule type: DNA	
A;Residues: 1-799 <WIL>	
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN: CAB03412.1	
A;Experimental source: clone T23F1	
A;Genetics:	
A;Gene: CESP:T23F1.7b	
A;Map position: 5	
A;Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1	

R.Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25173
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-779 <WLL>
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN: CAB03411.1
A:Experimental source: clone T23F1
C:Genetics:
A:Gene: CESP:T23F1.7a
A:Map position: 5
A:Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1
C:Superfamily: dipeptidyl-peptidase IV

Query Match 9.5%; Score 445.5; DB 2; Length 779;
Best Local Similarity 23.0%; Pred. No. 1.7e-23;
Matches 193; Conservative 126; Mismatches 292; Indels 229; Gaps 38;
QY 78 PHSDRIYLLAMS-----GENRENTLFYSEIPKTIINRAAVMLSMKPLDLFQATLDYGM 132
DB 116 PSADKRYFAMMDHAPMGMPQNETFHLKIVNNR-----ITYDIGL 158
QY 133 SREBELRERKRIGTVGIAASYDYGSGTFLFQAGSGIYHVKDGGPQGTQOPLRPNLVE 192
DB 159 -RKESV-----IQAFKWKGNFDFVFNKTIY----- 187
QY 193 TSCPNIRMDPKLCPADPDWIAFHSNDIWSIVTREERLLTYVHNELANMEEDARSAGV 252
DB 188 QSSP-----EEGLTRVNSGGEHVD-----GL 210
QY 253 ATFVLOEE-FDRYSGYWCPCAEETTPSGGKILRIIYEENDESEVEI IHVTSPMLETRAD 311
DB 211 FDIWYEEIEFGRDAMWSTK-----GDLAVASVDNHLTKVSL--KTVHLEPYPID 262
QY 312 -SPRYPT-----GTANPKVTFKSMEIMIDAGRIIDVIDKELQPPFIL-----FEGV 359
DB 263 TNFHYPTFAKVLPTVTLISWNKKE-----QSRQLDVLQKDSLSYHYLLAVKWLINGT 317
QY 360 EYIARAGWTPEGYKANSILL-----DRSOTRLQIVLISPELLFIPVEDDVMRQRLIESVP 414
DB 318 EQLVSV-WT--NRYQNEVALTICDWDTAICRLEFE----- 349
QY 415 DSVTPLIYYEETDIIWINIHIDIPHVPFQSGHEEIEFIFASECKTGFRHLKYITSILKESK 474
DB 350 -----YKASKRW-T-HDDFHSI-TSFDTLFELLP-----HDKEDNAFQQVAS 391
QY 475 YKRSGGLPAPSPKCPKEIEIAITSGEWEVLGRHGSNIQVDEVRRLVYFSGTKDSPLEH 534
DB 392 LRLSHGQLRTPK-----FLNLGEYDVTSINGINKET-----RTIFFHAAAPKPSHR 437
QY 535 HLYVSVYVNPGEVTRLTDRGYSHSCCIS--QHCOFFTSKYSNKQNPCHVSLYKLSPED 591
DB 438 SLFSYS-----LADESRSAYCISCSIKNCWAOQMDQMKTAIVSCKGPAPAPT 488
QY 592 -----DPTCKTK-----EFWATILDSAGPLPDVTPPISFESTTGTFLYGLMYK 636
DB 489 AIVNLTRMDSKTEHANILYDKTYQNRVEEAG-LPVIKETI-----KISDDFDALIKLSI 544
QY 637 PHDLQPGKKY---PTVLFIYGGPQVQLVNNRFGVKYFRLNTLASLGVVVVV--IDNRGSC 692
DB 545 PKDIYNRDKHQAIPLIVHVYGGNDQ---NTEATQIGIEEVVASQAAILIRIDGRGSG 601
QY 693 HRGLKPEGAFKYMGOIEIDDDVEGLQYLASRY-DFIDLDVRGIHGWYGGYLSMALMQ 751
DB 602 GRGWKYSAYIGOLGTVEVEDQIKAKVILRYRHLLDARRVAVFGWSYGGFMTLSMWNE 661
QY 752 RSD-1FRVAIAGAPVTLWIEYDGYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRL 810
DB 662 APEOFFKCAVSAPVTNFAYIDATYTERYWG--DAPLESY--SDVTKKLDNFKS--TELL 715
QY 811 LLHGFLDENVHFAHTSILLSFLVRACKPYDLOIYQPERHSIRVPESGHEYLHLHLVLOE 870

DB 716 LMHGLDDNVHVFQNSAILIDELQNRGVDFDLMMVFNQAHSLSSRTS--HVVGMKTHFLRQ 773
RESULT 13
A41793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41793
R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthold, R.J.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
A:Title: Differential expression of two distinct forms of mRNA encoding members of a di-
peptidyl-peptidase family
A:Reference number: A41793; MUID:92108018; PMID:1729689
A:Accession: A41793
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-803 <WAD>
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:9408719; PIDN:
A:Note: sequence extracted from NCBI backbone (NCBI:P:75138)
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptidase hydrolase; glycoprotein
F:257,342/Binding site: carbohydrate (Asn) #status predicted
Query Match 9.4%; Score 440; DB 2; Length 803;
Best Local Similarity 24.4%; Pred. No. 4.3e-23;
Matches 176; Conservative 114; Mismatches 317; Indels 114; Gaps 29;
QY 171 YHVKGPGQGTQOPLRPNLIVETSCPNIRMDPKLCPADPDWIAFHSNDIWSIVTREE 230
DB 156 YVLSKIPHGDPQSLQPPPEVSNAKLQYAGWGP-----GQLIFIFENNIIYCAHVQQA 210
QY 231 RRLTYVNEELANMEEDARSAGVATFVLOEE-FDRYSGYWCPCAEETTPSGGKILRIIYEE 289
DB 211 IRV-----VSTGKEGVIYNGLSMDWLYEBEILKTHIAHWSPDG-----TELAVAT 255
QY 290 NDESEVEI IHV---TSPMLETRADSFRYPKGTGTANPKVTFKMSMEIMIDAGRIIDVIDK 346
DB 256 INDSRVFVMLPTVYTGVSVPY--AKPYHYPKACGNCNPSISLH-----VIGLNGPTHDL--- 306
QY 347 ELIOPPEILFEGVEYIARAGWTPEGYK--WSILLDRSQTRLQIVLISPELLFIPVEDDVM 404
DB 307 EMTFPDPRREV-YITWKNWATSKVAVNW--LSRAQN-----VS----- 344
QY 405 ERQLIESVDSPTVPLII--YEETDIIWINIHIDIPHVPFQSGHEEIEFIFASECKTGPRH 462
DB 345 -----ILTLCDATTGVCTKKHDESEAWL-----HRQNEEPVFS---KDG-RK 383
QY 463 LYKITSILK--ESKYRSGGLPAPSDFKCPKEIEIAITSGEWEVLGRHGSNIQVDEVR 520
DB 384 FFFVRAIPQGGQGFYHITVSSSQPNSSNDNIQ---SITSGMDVT---KILSYDEKRS 436
QY 521 LVYFEGTKDPLEHLLVSVYVNPGEVTR-----LDRGYSHSCCISQHCDFIFISKYSN 574
DB 437 QIYFLSTEDLPRRQLYSASTV--GSFNQCLSCDLVNDCTYFSASFPGADFFLLKCEG 494
QY 575 QKNPHCVSLYKLSPEDDPTCKTKE-FWATILDSAGPLPDVTPPISFESTTGTFLYGM 633
DB 495 PGVP-TVSVHNTDKKMFDELNEHVQKAIASDRQMPKVEYRKIE-----TDDYNLPQ 547
QY 634 LYKPHDLQPGKKYPTVLFIYGGPQVQLVNNRFGVKYFRLNTLASLGVVVVVIDNRGSG 693
DB 548 ILKPEATETDAHYPLLLVVDGTGPGSQSVAEKF--AVTWETVMVSHGAVVCKDGRGSGF 605
QY 694 RGLAFEGAFKYMGOIEIDDDVEGLQYLASRYDFIDLDVRGIHGWYGGYLSMALMQRS 753
DB 606 QGTLLHLEVRRLRGLSEKQDMKQEAARVML-KEPYIDKTRVAVFGKDYGLSTYLLPAKG 664
QY 754 D-----1FRVAIAGAPVTLWIEYDGYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRL 809
DB 665 DQAPVFSCGSALSPITDFKLYASAFSERYLGLHGLDNRAVEMAKVHRVSAL--EGQOF 722
QY 810 LLHGFLDENVHFAHTSILLSFLVRACKPYDLOIYQPERHSIRVPESGHEYLHLHLVLO 869

Db 723 LVIHATADEKIHFOHTAELITQLIRKANKYSLQIYPDESHYFSAALQOHLHRSILGPFV 782
QY 870 E 870
Db 783 E 783
RESULT 14
168600
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: I68600
R:Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I68600
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-803 <RES>
A:Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:G306707; PID:AAA35761.1; PID:
C:Superfamily: dipeptidyl-peptidase IV
Query Match 9.1%; Score 427; DB 2; Length 803;
Best Local Similarity 23.4%; Pred. No. 3.7e-22;
Matches 172; Conservative 121; Mismatches 299; Indels 144; Gaps 30;
QY 171 YHVKDGGPQGFQQPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREE 230
Db 156 YVLSKIPHGDPQSLDPEVSNAKLQYAGWGPK-----GQQLIFIFENNIYYCAHVQKQA 210
QY 231 RRLTYVHNLANMEEDARSAGVATFVLOEE--FDYSGYWCPCAKETTPSGGKILRLIYEE 289
Db 211 IRV-----VSTGKEGVIYNGLSDWLYEEIILKTHIAHWMSPDG-----TRLAYAA 255
QY 290 NDESEVEI IHV---TSPMLETRRADSFYPTKGTANPKVTFPMSEIMIDAEGRIIDVIDK 346
Db 256 INDSRVPIMLPTTYSIYPT--VKPYHYPKAGSENPSISLH---VIGLNGPTHDL--- 306
QY 347 ELTQPEILLFEGVEYIARAGWTPEGKYA--WSILLDRSQRLQIVLISPELFIPEVDVDM 404
Db 307 EMPPDDPRMREY-YITMVKWATSTKVAVTW---LNRAQN-----VS----- 344
QY 405 ERORLESVPDSVTPLII--YEETDWINIHDPFVPPQSHREEI-----EFIFASEC 456
Db 345 -----ILTLCDATTGCTKKHEDESEAWLH-----RQNEBPVFSKDGKFFIRAI 390
QY 457 KTGFR-HLYKITSILKESYKSSGGLPAPSPDKPCPKKEEIAITSGGEWVLGRHGSNIQV 515
Db 391 PQGRGKGFYHIT--VSSQPNSSNDNIQ-----SITSGDMDVT---KILAY 431
QY 516 DEVRLVYFGTKDSPLEHLLYVSVNPGEVTRLTDRGYSH---SCCISOHCDFFIISKY 572
Db 432 DEGNKIYFLSTEDLPRRLQYSAN-----TEGNFNRQCLSCDLVENCITYFSASF 481
QY 573 SNQKNPHCVSLYKLSGP-----EDDPTCKTEF-----WATILDSAGPLPDYTPPE 618
Db 482 SHSMD---FELLKCEGPGVPMVTVHTTDDKKQMFLETNEHVKKAINDRQMPKVEYRDIE 538
QY 619 IFSPESTGTFTLYGMLYKPHDLQPKKKYPTVLFYGGPQVQLVNNRFGVKYFRLNLTAS 678
Db 539 I-----DDYNLPMQLIKPATFTDTHYPLLLVWDGTFGSGSVAEKFE--VSMETVMVSS 590
QY 679 LGYVVVVVIDNRGSHRGKLPFGAFKYMKGQIIEIDQVGLQVYLASRYDFIDLDVRGIGHW 738
Db 591 HGAVVVKCDGRGSGFGQTKLLHEVRRRLGLLEEKQMEAVRTML--KEQYIDRTVAVFGK 649
QY 739 SYGGLYSIMALMORSD-----IFRVAITAGAPVTLWIFDYDTGYTERVMGHPDQNEQGYILGS 794
Db 650 DYGGYLTSTYILPAKENQGTFTCGSALSPIITDFDKLYASAFSERVYLGLHGLDNRAYEMTK 709
QY 795 VMAQAEKPEPSEPNLLLLHGFIDENVHFPAHTSILLSSFLVRACKPYDLQIYQERHSIRVP 854

Db 710 VAHRVSAL--EEQQFLIHIPTADEKIHFOHTAELITQLIRKANKYSLQIYPDESHYFTSS 767
QY 855 ESGEHYELHLHYLOE 870
Db 768 SLKQHLRYRSIINFFVE 783
RESULT 15
I54331
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I54331
R:Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I54331
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-865 <RES>
A:Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:G306705; PIDN:
C:Superfamily: dipeptidyl-peptidase IV
Query Match 9.1%; Score 427; DB 2; Length 865;
Best Local Similarity 23.4%; Pred. No. 4.1e-22;
Matches 172; Conservative 121; Mismatches 299; Indels 144; Gaps 30;
QY 171 YHVKDGGPQGFQQPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREE 230
Db 218 YVLSKIPHGDPQSLDPEVSNAKLQYAGWGPK-----GQQLIFIFENNIYYCAHVQKQA 272
QY 231 RRLTYVHNLANMEEDARSAGVATFVLOEE--FDYSGYWCPCAKETTPSGGKILRLIYEE 289
Db 273 IRV-----VSTGKEGVIYNGLSDWLYEEIILKTHIAHWMSPDG-----TRLAYAA 317
QY 290 NDESEVEI IHV---TSPMLETRRADSFYPTKGTANPKVTFPMSEIMIDAEGRIIDVIDK 346
Db 318 INDSRVPIMLPTTYSIYPT--VKPYHYPKAGSENPSISLH---VIGLNGPTHDL--- 368
QY 347 ELTQPEILLFEGVEYIARAGWTPEGKYA--WSILLDRSQRLQIVLISPELFIPEVDVDM 404
Db 369 EMPPDDPRMREY-YITMVKWATSTKVAVTW---LNRAQN-----VS----- 406
QY 405 ERORLESVPDSVTPLII--YEETDWINIHDPFVPPQSHREEI-----EFIFASEC 456
Db 407 -----ILTLCDATTGCTKKHEDESEAWLH-----RQNEBPVFSKDGKFFIRAI 452
QY 457 KTGFR-HLYKITSILKESYKSSGGLPAPSPDKPCPKKEEIAITSGGEWVLGRHGSNIQV 515
Db 453 PQGRGKGFYHIT--VSSQPNSSNDNIQ-----SITSGDMDVT---KILAY 493
QY 516 DEVRLVYFGTKDSPLEHLLYVSVNPGEVTRLTDRGYSH---SCCISOHCDFFIISKY 572
Db 494 DEGNKIYFLSTEDLPRRLQYSAN-----TEGNFNRQCLSCDLVENCITYFSASF 543
QY 573 SNQKNPHCVSLYKLSGP-----EDDPTCKTEF-----WATILDSAGPLPDYTPPE 618
Db 544 SHSMD---FELLKCEGPGVPMVTVHTTDDKKQMFLETNEHVKKAINDRQMPKVEYRDIE 600
QY 619 IFSPESTGTFTLYGMLYKPHDLQPKKKYPTVLFYGGPQVQLVNNRFGVKYFRLNLTAS 678
Db 601 I-----DDYNLPMQLIKPATFTDTHYPLLLVWDGTFGSGSVAEKFE--VSMETVMVSS 652
QY 679 LGYVVVVVIDNRGSHRGKLPFGAFKYMKGQIIEIDQVGLQVYLASRYDFIDLDVRGIGHW 738
Db 653 HGAVVVKCDGRGSGFGQTKLLHEVRRRLGLLEEKQMEAVRTML--KEQYIDRTVAVFGK 711
QY 739 SYGGLYSIMALMORSD-----IFRVAITAGAPVTLWIFDYDTGYTERVMGHPDQNEQGYILGS 794
Db 712 DYGGYLTSTYILPAKENQGTFTCGSALSPIITDFDKLYASAFSERVYLGLHGLDNRAYEMTK 771

Qy 795 VAMQAEKFFSEPNRLLLHGFLDENVHFAHTSILLFLVRAGKPYDLOIYPQERHSIRVP 854
Db 772 VAHRVSAL--EEQFLIIHPTADEKIHQHTAELITQLIRGKANYSLQIYPDESHYFTSS 829
Qy 855 ESCEHYELHLLHYLQE 870
Db 830 SLKQHLYSIINFVE 845

Search completed: April 14, 2006, 12:51:03
Job time : 55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 12:43:37 ; Search time 240 Seconds
(without alignments)
2592.817 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMETQLGVETADCT.....HLLHLYQENLGRIRALKVI 882

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Databases: UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4700	100.0	898	1	DPB8_HUMAN	Q6V1X1 homo sapien
2	4336.5	96.5	892	1	DPB8_MOUSE	Q80Y47 mus musculus
3	2870	61.1	863	1	DPB9_HUMAN	Q86T12 homo sapien
4	2833	60.3	862	1	DPB9_MOUSE	Q8BV34 mus musculus
5	2829.5	60.2	923	2	Q4SBM6_TETNG	Q4SBM6 tetraodon n
6	2707.5	57.6	847	2	Q6GR22_XENLA	Q6GR22 xenopus lae
7	1643.5	35.0	508	2	Q75273_HUMAN	Q75273 homo sapien
8	1641	34.9	886	2	Q70BK1_ANOGA	Q70BK1 anopheles g
9	1596	34.0	1053	2	Q9VC20_DROME	Q9VC20 drosophila
10	1596	34.0	1113	2	Q9VC19_DROME	Q9VC19 drosophila
11	1567	33.3	740	2	Q5TTK8_ANOGA	Q5TTK8 anopheles g
12	1321	28.1	621	2	Q7PTT8_ANOGA	Q7PTT8 anopheles g
13	1142.5	24.3	557	2	Q5TXJ2_ANOGA	Q5TXJ2 anopheles g
14	1124.5	23.9	803	2	Q54U01_DICTDI	Q54U01 dictyosteli
15	1049	22.3	432	2	Q75868_HUMAN	Q75868 homo sapien
16	924.5	19.7	927	2	Q965K3_CAEL	Q965K3 caenorhabdi
17	922.5	19.6	931	2	Q44987_CAEL	Q44987 caenorhabdi
18	867.5	18.5	895	2	Q61CU7_CAEBR	Q61CU7 shewanella
19	846	18.0	763	2	Q8EAB7_SHEON	Q8EAB7 shewanella
20	842	17.9	738	2	Q9A6E0_CAUCR	Q9A6E0 caulobacter
21	837.5	17.8	746	2	Q9FNF6_ARATH	Q9FNF6 arabidopsis
22	792.5	16.9	745	2	Q6F317_9PSED	Q6F317 pseudomonas
23	789	16.7	596	2	Q6K880_ORISA	Q6K880 oryza sativ
24	779	16.6	743	2	Q5QX36_IDILO	Q5QX36 idiomarina
25	759	16.1	745	2	Q5H5W8_XANON	Q5H5W8 xanthomonas
26	754	16.0	741	2	Q95782_XANNA	Q95782 xanthomonas
27	751.5	16.0	751	2	Q4UPD3_XANCP	Q4UPD3 xanthomonas
28	751.5	16.0	751	2	Q8P3V8_XANCP	Q8P3V8 xanthomonas
29	746	15.9	766	2	Q4TNP1_9SPHN	Q4TNP1 erythrobact
30	744.5	15.8	757	2	Q8PFD7_XANAC	Q8PFD7 xanthomonas
31	727	15.5	552	2	Q8GUJ7_ARATH	Q8GUJ7 arabidopsis

32	724.5	15.4	720	2	Q5NMW8_ZYMMO	Q5NMW8 zymomonas m
33	718.5	15.3	749	2	Q7NEK8_GLOVI	Q7NEK8 gloeobacter
34	672	14.3	736	2	Q8A028_BACTN	Q8A028 bacteroides
35	669.5	14.2	739	2	Q5FQY6_GLUOX	Q5FQY6 gluconobact
36	655.5	13.9	850	2	Q6H9E3_TRYP	Q6H9E3 trypanosoma
37	651	13.9	736	2	Q5LGU5_BACFN	Q5LGU5 bacteroides
38	648	13.8	736	2	Q64XP9_BACFR	Q64XP9 bacteroides
39	643	13.7	852	2	Q4QIH9_LEIMA	Q4QIH9 leishmania
40	620	13.2	711	2	Q47900_FLAME	Q47900 flavobacter
41	595.5	12.7	778	2	Q51ZP7_MAGGR	Q51ZP7 magnaporthe
42	587	12.5	769	2	Q5QVY7_IDILO	Q5QVY7 idiomarina
43	587	12.5	809	2	Q4NVV1_9DELT	Q4NVV1 anaeromyxob
44	582.5	12.4	771	2	Q42812_ASPOR	Q42812 aspergillus
45	578	12.3	731	2	Q6L872_PREIN	Q6L872 prevotella

ALIGNMENTS

RESULT 1
ID DPB8_HUMAN STANDARD; PRT; 898 AA.
AC Q6V1X1; Q7Z4C8; Q7Z4D3; Q7Z4E1; Q81WG7; Q8NEM5; Q96JX1; Q9HEM2;
AC Q9HEM3; Q9HEM4; Q9HEM5; Q9NXP4;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DE Dipterydyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DP8)
DE (Prolyl dipeptidase DPP8) (Dipeptidyl peptidase IV-related protein 1) (DPP8-1).
DE Name=DPP8; Synonyms=DPBP1; ORFNames=MSTP097, MSTP135, MSTP141;
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION, CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION, AND SUBCELLULAR LOCATION.
RC TISSUE=Placenta;
RX MEDLINE=20467194; PubMed=11012666;
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W., Gorrell M.D.;
RT "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8";
RL Eur. J. Biochem. 267:6140-6150(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME REGULATION, BIOPHYSICOCHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Testis;
RX PubMed=12662155; DOI=10.1042/BJ200021914;
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinsanya K.O.;
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases";
RL Biochem. J. 373:179-189(2003).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
RC TISSUE=Testis;
RA Sha J.H., Zhou Z.M., Li J.M.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 211-898 (ISOFORM 2).
RC TISSUE=Hepatooma, and Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Ohtsuka M., Ishii T., Shiba T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,

RA Nagahari K., Murakami K., Yasuda T., Iwavanagi T., Wagatsuma M.,
RA Shitatori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Onura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hota T.,
RA Kusano J., Tangehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Koguchi S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Inose N., Mueashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Koniya M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [5].
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC MEDLINE:22388257; PubMed:12477932; DOI=10.1073/pnas.242603899;
EX Strausberg R.L., Feingold E.A., Grouse L.H., Dergos J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebrow K.H., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Bhat S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RC [6].
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 561-898.
RC TISSUE=Arteria;
RA Zhao B., Xu H.S., Tong Y.K., Sheng H., Qin B.M., Liu Y.Q., Liu B.,
RA Wang X.Y., Zhang Q., Song L., Gao Y., Zhang C.L., Ye J., Ji X.J.,
RA Liu B.H., Lu H., Chen J.Z., Cai M.Q., Zheng W.Y., Teng C.Y., Liu Q.,
RA Yu L.T., Lin J., Gong Q., Zhang A.M., Gao B.L., Rui R.T.,
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [7].
RP MUTAGENESIS OF GLU-275; SER-755; ASP-833 AND HIS-865. CATALYTIC
RP ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, AND SUBCELLULAR LOCATION.
EX PubMed:12534281; DOI=10.1021/bi026846s;
RA Ajami K., Abbott C.A., Obradovic M., Gyabers V., Kaehne T.,
RA McCaughan G.W., Gorrell M.D.,
RT "Structural requirements for catalysis, expression, and dimerization
RL in the CD26/DPPIV gene family.";
RL Biochemistry 42:694-701(2003).
RN [8].
RP BIOPHYSICOCHEMICAL PROPERTIES.
EX PubMed:15039077; DOI=10.1016/j.pep.2003.12.019;
RA Chen Y.-S., Chien C.-H., Goparaju C.M., Hsu J.T.-A., Liang P.-H.,
RA Chen X.;

"Purification and characterization of human prolyl dipeptidase DPP8 in
Sf9 insect cells";
Protein Expr. Purif. 35:142-146(2004).
RN [9].
RP ENZYME REGULATION
EX PubMed:1566483; DOI=10.1016/j.bmc.2004.11.023;
RA Chang W.-T., Chen Y.-S., Hsu T., Wu S.-H., Chien C.-H., Chang C.-N.,
RA Chang S.-P., Lee S.-J., Chen X.;
RT "Novel isoindoline compounds for potent and selective inhibition of
prolyl dipeptidase DPP8.";
RL Bioorg. Med. Chem. Lett. 15:687-691(2005).
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
dipeptides from proteins having a Pro or Ala residue at position
2. May play a role in T-cell activation and immune function.
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
Zaa is neither Pro nor hydroxyproline.
CC -!- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
protease inhibitor 4-(2-aminoethyl)benzenesulfonyl fluoride
(AESBF), and by di-isopropylfluorophosphate. Specifically inhibited
by isoindoline derivatives.
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
Kinetic parameters:
KM=208 uM for Ala-Pro-AMC;
KM=130 uM for Ala-Pro-AFC;
KM=120 uM for H-Ala-Pro-pNa;
KM=1420 uM for H-Ala-Ala-pNa;
KM=310 uM for H-Arg-Pro-pNa;
KM=2050 uM for H-Asp-Pro-pNa;
KM=480 uM for H-Gly-Pro-pNa;
pH dependence:
Optimum pH is 7.4-8.5. Little activity below pH 6.5;
-!- SUBCELLULAR LOCATION: Cytoplasmic.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=6;
Name=1;
IsoId=Q6VLX1-1; Sequence=Displayed;
Name=2;
IsoId=Q6VLX1-2; Sequence=VSP_013864;
Name=3;
IsoId=Q6VLX1-3; Sequence=VSP_013860;
Name=4;
IsoId=Q6VLX1-4; Sequence=VSP_013860, VSP_013862;
Name=5;
IsoId=Q6VLX1-5; Sequence=VSP_013863;
Name=6;
IsoId=Q6VLX1-6; Sequence=VSP_013861;
-!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
testis, placenta, prostate, muscle and brain.
-!- INDUCTION: In activated T-cells.
-!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
-!- CAUTION: Ref.4 (BAA91059) sequence differs from that shown due to
a frameshift in position 486.
-!- CAUTION: Ref.6 (AAQ13650 and AAQ13623) sequences differ from that
shown due to several frameshifts.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; AF221634; AAG29766.1; -; mRNA.
EMBL; AF221635; AAG29767.1; -; mRNA.
EMBL; AF221636; AAG29768.1; -; mRNA.
EMBL; AF221637; AAG29769.1; -; mRNA.
EMBL; AF172859; AAG17261.1; -; mRNA.
EMBL; AF172859; AAG17261.1; -; mRNA.
EMBL; AF354302; AAG63887.1; -; mRNA.
EMBL; AK000290; BAA91059.1; ALT_FRAME, mRNA.
EMBL; AK027826; BAB55395.1; ALT_INIT, mRNA.
EMBL; BC030888; AAB30888.3; -; mRNA.
EMBL; BC040203; AAB40203.1; ALT_INIT, mRNA.
EMBL; AF176779; AAG13657.1; ALT_INIT, mRNA.

RL Mol. Cell. Proteomics 3:279-286(2004).
CC - FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position
CC 2. May play a role in T-cell activation and immune function (By
CC similarity).
CC
CC - CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa- from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC - ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
CC proteinase inhibitor 4-(2-aminocetyl)benzenesulphonyl fluoride
CC (ABSPF), and by di-isopropyl fluorophosphate. Specifically inhibited
CC by isodoline derivatives (By similarity).
CC - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC - SIMILARITY: Belongs to the peptidase S9B family. peptiv subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AK016546; BAB30295.2; -; mRNA.
CC EMBL; BC043124; AAH43124.1; -; mRNA.
CC EMBL; BC059222; AAH59222.1; -; mRNA.
CC MEROPS; S09.018; -;
CC Ensembl; ENSMUSG0000032393; Mus musculus.
CC
CC MGI; MGI:1921638; Dpp8.
CC InterPro; IPR001375; Peptidase S9.
CC InterPro; IPR002469; Peptidase S9B.
CC InterPro; IPR000379; Ser esters.
CC Pfam; PF00930; DppIV_N; 1.
CC Pfam; PF00326; Peptidase S9; 1.
CC Asinopeptidase; Hydrolase; Phosphorylation; Protease; Serine protease.
KW ACT_SITE 749 749 Charge relay system (By similarity).
FT ACT_SITE 827 827 Charge relay system (By similarity).
FT ACT_SITE 859 859 Charge relay system (By similarity).
FT ACT_SITE 325 325 Phosphotyrosine.
FT MOD_RES 87 87 G - (in Ref. 1).
FT CONFLICT 892 AA; 102186 MW; 59081CD9792E03ED CRC64;
SQ SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;

Query Match 96.5%; Score 4536.5; DB 1; Length 892;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 846; Conservative 18; Mismatches 18; Indels 1; Gaps 1;

Qy 1 MAAMATEQLGVIEFETADCEE-NIESQDRKLEPFYVERYSWOLKLLADTRKYGVM 59
Db 10 MAAMATEQLGVIEFETAECEGNGESQDRKLEPFYVERYSWOLKLLADTRKYGVM 69

Qy 60 MAKAPHDFMFKRNDPDGPHSDRIYILAMSGENRENTLFYSEIPKTNRAAVLMWKL 119
Db 70 MAKAPHDFMFKRNDPDGPHSDRVYLLAMSGENRENTLFYSEIPKTNRAAVLMWKL 129

Qy 120 LDLFQATLDYGMYSREELLRERKRGITGVGTASVDYHOGSGTFLFQAGSGIYHVKGQP 179
Db 130 LDLFQATLDYGMYSREELLRERKRGITGVGTAAVDYHFGSGTFLFQAGSGIYHVKGQP 189

Qy 180 GFTQOPLRNLVTSFCNRMIPDKLPADPDWIAFHSNDIWIISNIVTREERLTYVHNE 239
Db 190 GFTQOPLRNLVTSFCNRMIPDKLPADPDWIAFHSNDIWIISNIVTREERLTYVHNE 249

Qy 240 LANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGKILRIILYENDESEVBIH 299
Db 250 LANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGKILRIILYENDESEVBIH 309

Qy 300 VTSPLMTRADSFYKPKGTANPKVTFKMSIMIDARGRIIDVIDKELQPFILPEGV 359
Db 310 VTSPLMTRADSFYKPKGTANPKVTFKMSIIVDAAGIIDVIDKELQPFILPEGV 369

Qy 360 EYIARAGWTPEGKANSILLDRSQTLQIVLISPELFTPVDDVYMERORLIESVDSVTP 419
Db 370 EYIARAGWTPEGKANSILLDRSQTLQIVLISPELFTPVDDVYMERORLIESVDSVTP 429

Qy 420 LIIYEETTDIWINIHDFHVPFQTHSEETFEFASECKTGFPHLYKITSILKESYKRSS 479

Db 430 LIIYEETTDIWINIHDFHVPFQTHSEETFEFASECKTGFPHLYKITSILKESYKRSS 489
Qy 480 GGLPAPSDFKPIKEETAITSGEVLGRHSGNIOWDEVRVLYFEGTKDSPLEHLVYV 539
Db 490 GGLPAPSDFKPIKEETAITSGEVLGRHSGNIOWDEARLKYFEGTKDSPLEHLVYV 549
Qy 540 SYVNPGEVRLTRDGYSHSCCISCHOCFFISKYSNQNPHCVSLYKLSPPDDPTCKTKE 599
Db 550 SYVNPGEVRLTRDGYSHSCCISCHOCFFISKYSNQNPHCVSLYKLSPPDDPTCKTKE 609
Qy 600 FWATILDSAGLPDYTPPEIPSPFESTTGTFLYGMVYKPHDLOPGKKYPTVLVIYGGPQV 659
Db 610 FWATILDSAGLPDYTPPEIPSPFESTTGTFLYGMVYKPHDLOPGKKYPTVLVIYGGPQV 669
Qy 660 LVNNRFKGVKVFRLNTLASGLYVVYVVDNRGSGHRLKPEGAFKYMGOIEIDDDVEGLQ 719
Db 670 LVNNRFKGVKVFRLNTLASGLYVVYVVDNRGSGHRLKPEGAFKYMGOIEIDDDVEGLQ 729
Qy 720 YLASRYDFIDLDVRVGIHGSYGGYLSLMLMQRSDFRVAIAGAPVTLWIFYDTGYTERY 779
Db 730 YLASRYDFIDLDVRVGIHGSYGGYLSLMLMQRSDFRVAIAGAPVTLWIFYDTGYTERY 789
Qy 780 MGHPPDQSGYVLSGVAMQAEKPPSEPNRLLLHGLFDENVHFAHTSILLSLVLRAGKY 839
Db 790 MGHPPDQSGYVLSGVAMQAEKPPSEPNRLLLHGLFDENVHFAHTSILLSLVLRAGKY 849
Qy 840 DLQIYQERHSIRVPESGEHVELHLLHYLOENLGSRIAALKVI 882
Db 850 DLQIYQERHSIRVPESGEHVELHLLHYLOENLGSRIAALKVI 892

RESULT 3
DPP9 HUMAN STANDARD; PRT; 863 AA.
ID DPP9 HUMAN Q6A137; Q6A137; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WXD8;
AC Q6NT8; Q9BVR3;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 9 (EC 3.14.5) (Dipeptidyl peptidase IX) (DPP9)
DE (Dipeptidyl peptidase-like protein 9) (DPLP9) (Dipeptidyl peptidase
DE IV-related protein 2) (DPRP-2).
DE Name=DPP9; Synonyms=DPRP2;
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.
RX MEDLINE=22347328; PubMed=12459266; DOI=10.1016/S0378-1119(02)01059-4;
RA Olsen C., Wagtmann N.;
RT "Identification and characterization of human DPP9, a novel homologue
RT of dipeptidyl peptidase IV.";
RL Gene 299:185-193(2002).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), CATALYTIC ACTIVITY,
RP BIOPHYSICO-CHEMICAL PROPERTIES, ENZYME REGULATION, TISSUE SPECIFICITY,
RP AND SUBCELLULAR LOCATION.
RC TISSUE=Colon;
RX PubMed=12662155; DOI=10.1042/BJ20021914;
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akineanya K.O.;
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member
RT of an emerging subgroup of serine proteases.";
RL Biochem. J. 373:179-189(2003).
[3]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), PARTIAL NUCLEOTIDE SEQUENCE
RP [MRNA] (ISOFORM 2), CATALYTIC ACTIVITY, BIOPHYSICO-CHEMICAL PROPERTIES,
RP TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX PubMed=15245913; DOI=10.1016/j.bbexp.2004.03.010;
RA Ajami K., Abbott C.A., McCaughan G.W., Gorrell M.D.;
RT "Dipeptidyl peptidase 9 has two forms, a broad tissue distribution,

RT cytoplasmic localization and DPV-like peptidase activity.";
 RL Biochim. Biophys. Acta 1679:18-28(2004).
 RN [4].
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
 RC TISSUE=Placenta, and Skin;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapellato M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5].
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 30-863 (ISOFORM 3),
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 272-863 (ISOFORM 2), AND
 RP PARTIAL NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 5).
 RC TISSUE=Glial tumor, Ovary, Spleen, and Trachea;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ora T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
 RA Tshida S., Ono Y., Takiguchi S., Watanabe M., Yosida M., Hota T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Inoue N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,
 RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isoigai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [6].
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 209-863 (ISOFORM 4), AND
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 298-863 (ISOFORM 2).
 RC TISSUE=Melanoma;
 RG The German cDNA consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
 CC dipeptides from proteins having a Pro or Ala residue at position
 CC 2.
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa|-
 Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
 Zaa is neither Pro nor hydroxyproline.
 CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
 CC (2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-
 CC isopropylfluorophosphate.
 CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
 CC Kinetic parameters:
 CC KM=161 uM for Ala-Pro-AMC;
 CC KM=180 uM for Ala-Pro-AFC;
 CC pH dependence:
 CC Optimum pH is 7.5-8.5. Little activity below pH 6.5;
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; cytosol.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=5;
 CC Name=1; Synonyms=Short;
 CC IsoId=Q86TI2-1; Sequence=Displayed;
 CC Name=2; Synonyms=Long;
 CC IsoId=Q86TI2-2; Sequence=VSP_013865;
 CC Note=Incomplete sequence;
 CC Name=3;
 CC IsoId=Q86TI2-3; Sequence=VSP_013867, VSP_013868;
 CC Name=4;
 CC IsoId=Q86TI2-4; Sequence=VSP_013869;
 CC Name=5;
 CC IsoId=Q86TI2-5; Sequence=VSP_013865, VSP_013866;
 CC Note=Incomplete sequence;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
 CC liver, heart and muscle, and lowest levels in brain.
 CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC -!- CAUTION: Ref.6 (CAD39039) sequence differs from that shown due to
 CC frameshifts in positions 432 and 460.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: AF452102; AAL47179.1; -; mRNA.
 CC EMBL: AY172660; AAO17262.1; -; mRNA.
 CC EMBL: AF542510; AAO73880.2; -; mRNA.
 CC EMBL: AY374518; AAQ83119.1; -; mRNA.
 CC EMBL: BC000970; AAH00970.1; -; mRNA.
 CC EMBL: BC037948; AAH37948.1; -; mRNA.
 CC EMBL: AK054656; BAB70784.1; ALT_INIT; mRNA.
 CC EMBL: AK075030; BAC1362.1; -; mRNA.
 CC EMBL: AK131499; BAC85150.1; -; mRNA.
 CC EMBL: AK131499; BAD18643.1; ALT_INIT; mRNA.
 CC EMBL: AL834376; CAD39039.3; ALT_FRAME; mRNA.
 CC EMBL: CR627380; CAH10477.1; -; mRNA.
 CC MEROPS: S09.019; -;
 CC Ensembl: ENSG00000142002; Homo sapiens.
 CC HGNC: HGNC:18648; DPP9.
 CC MIM: 608258; -;
 CC InterPro: IPR001375; Peptidase_S9.
 CC InterPro: IPR002469; Peptidase_S9B.
 CC InterPro: IPR000379; Ser_estrs.
 CC Pfam: PF00930; DPPIV_N.1.
 CC Pfam: PF00326; Peptidase_S9; 1.
 CC Alternative splicing; Amino-peptidase; Hydrolase; Protease;
 CC Serine protease.
 CC ACT_SITE 730 730 Charge relay system (By similarity).
 CC ACT_SITE 808 808 Charge relay system (By similarity).
 CC ACT_SITE 840 840 Charge relay system (By similarity).
 CC VARSPPLIC 1 1 M -> LSRRVPCVRGCRPPPLPPLPGSOSRAWSDREAPLD
 CC PGRPAQSGRRPTSRSGACSWNGSGSLDPLEGTALURSAE
 CC RLMRKVKLRDKENTGWSRFSLSNESAER (in
 CC isoform 2 and isoform 5).
 CC /FTId=VSP_013865.
 CC Missing (in isoform 5).
 CC /FTId=VSP_013866.
 CC Missing (in isoform 5).
 CC VARSPPLIC 650 674 QLVNNSFGIKYLRINTLASLGYAV -> SAHLLPRPPPHH
 CC PPDSPSLKQCL (in isoform 3).
 CC FT

Tue Apr 18 08:18:35 2006

RA Nicad S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cottolier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardit G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter E.S., Weissenbach J., Roest Crollius H.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAEE01014667; CAG01956.1; -; Genomic_DNA.
DR NON_TER 1
FT SEQUENCE 923 AA; 105211 MW; 1397023B2004D009 CRC64;
SQ
Query Match 60.2%; Score 2829.5; DB 2; Length 923;
Best Local Similarity 56.6%; Pred. No. 3.6e-197;
Matches 524; Conservative 138; Mismatches 190; Indels 73; Gaps 7;
19 DCENIESQDRPKLPFFVYVYSWQLKKLLADTRKYHYGYMAKAPHDFFVKNRDPDG 78
1 DSTEVEVMEDEVP--SQFFVQKHSWEGLRDIHCSCRKNSGIANKAPHDFQFVKQDENG 58
79 HSDRIYVLAMSGENRENTLFSEIPKTNRAAVLMLSWKPLLDLFOATLDYGNYSREEL 138
59 HSHRLYLGLMPYSGRNSLLYSEIPKVRKEALLVLSWKQMLDHFQATPHQGAYSREEL 118
139 LRERKRGITGLASVDYHGGSTFLFQAGSGYVHVGQGGFQFQQ--PLRPNLVETSCPN 197
119 LRERKLGAGITSYDYHAQTGLFLFQASNSLFYQDGGQNSFQSAPVXPKVEIKTQCSG 178
198 IRMDPKLPADPDWTAFTHSNDIWTLSNITVREERLTVYNELANMEEDARSAGVATFVL 257
179 TRMDPKLPANPDPIAFINNDLWVANIKTGEERLTFCCHKGADSVKEDPKSAGVATFVI 238
258 QEBEDRYSGVWCPKAEITPSSGKILRLIYENDESEVEIIVHTSPMLETRRADSPRYPK 317
239 QEBFDRTGTWSPSAVEDPDGQKRVLLYEVDETEVEIIVHVPSPALEERKADAVYPR 298
318 TGTANPKVTPKMSSEIMDAEGR-----IIDVIDKELIQP 351
299 TGSKNPQATIKLVEIKTDQGRVSLCRLLCVFLLRSDRAVLTCYLQIVSTQDKELAVP 358
352 FEILFEGVEYIARAGWTPGKYAWSTLLDRSOTRLOIVLISPFLPIVEDDVMERQELIE 411
359 FTSLFPCTEYIARVGTSGDKYGWAALLDRSOKQLVLLPFAFPVTDPPARQESLE 418
412 SVDPSTPLIYETTDIMINHDIPFVPPQSHHEIEPIFASECKTGFRHLKYITSILK 471
419 AVPDRTQYIYETTDVWNVHDIFYPFVQTAEDFTFIWNESKGTGFSHLKYITSVLH 478
472 -----ESKYRSGG-----LPA-----PSDFKCPKEIATSGEVEVLGRHSIQV 515
479 PGFHCWAEVHYHTGEDPQRIPAVSTDPVPGDFKCAVKEEITLTSGEVEVLARHGSKIW 538
516 DEVRLLVYFEGTKDPSLEHLLYVVSYNVPEVTRTLDRGYSHSCCIS----- 562
539 NESSKLVYFQGTROTPLLEHLLYVVSVDSPGVVRLTKFGFSHSCSVQKKSLOSDFYFN 598
563 -----QHCDFFIISKYQKXNPHCVSLYKLSSPDPTCKTKTEFWATILDS 607
599 YSSITLPLSLSSFIWQNFDFVSHVSVCTPCPCVHYVYKLNSSESDPLHIVPEFWASMMES 658

QY 608 AGPLPDYTPPPIPSFESTTGTFTLYGMLYKPHDLQPGKYPYTVLTYGSPQVOLVNNRFGK 667
DB 659 SGCPGDSVPPPIFDFOQKSGFYGMWYKPHSLQPGKHTVLFYVGSPQVOLVNNRFGK 718
QY 668 VKYFRNLNTASLGIVVVVVIINRNSCHRGKLPFGAFKFKYKGOLEIDDOVEGLQYLASRYDF 727
DB 719 MKYLRNLNTASLGIVVVVVIDGRSCORGLFEGSALKNKGQVEIDDOVEGLQYVAEKFN 778
QY 728 IDLDVRVGHWSYGYLSLMLMQRSDIFRVAIAGAPVTLWTFYDTGYTERYMGHPDQNE 787
DB 779 VDLRSVAIHGWSYGGFLSLMGLIQRPNVFKLAIAGAPVTVMWYDGTGYTERYMDVPENNQ 838
QY 788 QGYLYGSVAMQAEKFPSEPNRLLLHGFLENVHFHFAHTSLLSFLVRAKPKYDQLYQPOR 847
DB 839 QGYEESGVALHVDKLPSEPNRLLLHGFLENVHFHFAHTSLLSFLVRAKPKYDQLYQVNE 898
QY 848 RHSIRVPESGEHYELHLLHVLQENL 872
DB 899 RHSIRCPESGEHYELHLLHVLQOYL 923
RESULT 6
O6GR22 XENLA PRELIMINARY; PRT; 847 AA.
ID O6GR22; AC
DT 05-JUL-2004 (tREMBLrel. 27, Created)
DT 05-JUL-2004 (tREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (tREMBLrel. 27, Last annotation update)
DE MGC81313 protein.
GN Names=MGC81313
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=223388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berger J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Pringle C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullay J.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Small M.A.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX Klein S., Gerhard D.S.,
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.

[illegible]

Tue Apr 18 08:18:35 2006

```
Db 386 AIAGAPVTWVAYDTGYTERYMDVPENNOHQYBAGSVALLHVEKLPNEPRLILLHGFIDE 445
QY 819 NVHFAHTSILLSPVLRAGKPYDL-----OYPOERHSIRVPESGEHYELHLLHYIQ 869
Db 446 NVHFPHTNFLVSQLIRAGKPVQLVALPPVSPQIYPNERHSIRCPESGEHYEVTLLHFLQ 505
QY 870 ENL 872
Db 506 EYL 508

RESULT 8
ID Q7QBKI_ANOGA PRELIMINARY; PRT; 886 AA.
AC Q7QBKI;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000016526 (Fragment).
GN ORFNames=ENSANGG0000014037;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RT "The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RT "The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA0801008879; EAA08416.2; -; Genomic_DNA.
DR MEMOPS; S091016; F:catalytic activity; IEA.
DR GO; GO:0003824; F:kinetic-type peptidase activity; IEA.
DR GO; GO:0008236; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_Nr.1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1
FT NON_TER 886
SQ SEQUENCE 886 AA; 100440 MW; 3C284605CAA57DB4 CRC64;

Query Match 34.9%; Score 1641; DB 2; Length 886;
Best Local Similarity 39.2%; Pred. No. 1.8e-110;
Matches 361; Conservative 148; Mismatches 287; Indels 124; Gaps 20;

QY 41 SWSLKKLLADTRKYHYGMMAKA-PHDFMFVKNRDPDGHSDRIYYLAMSGENRNT-LF 98
Db 2 SWSLKKLVNTRQMATTAASFPMSVNF--RTLSDG--RTRVYFLSPPPANGWDITLF 57
QY 99 YSEIPKTIINRAAVLM-----LSWKPLLD 121
Db 58 YADVPPAGAKRSQLLPALDANEERLEADGSDGDEDAEGERRSAGRKHLFPWQLE 117
QY 122 LFOATLDYGVSRREELLRKRKIGTVGTASYDYHQSGTFLQAGSGIYHKVGGPQGF 181
Db 118 SVLGHLSST-NSREVQMLERKRLSWGTSVELHKAAGKIVFPACNTLYQCLD---TG 173
QY 182 TQOPLRNLVETSCPNIRMPDKLPADPDWIAFIHSDNIWISNIVTREERLTIVVHNELA 241
Db 174 EENPLFSELRILQRAAIDPQICPQNSDLVAPVCNGDIWVHTHSGHSERLTVAHDGR 233
```

```
QY 242 NNEEDARSAGVATVLOEEFDRIYSGYWCPCAKETTPSGGKILRIILYEENDESEVLIHVT 301
Db 234 SFAEDPLTAGVPSTVMQEEFESRIQGFWMQPE-----SHDEVVRIYVEEDESVDLYTFP 288
QY 302 SPMLETRADSRFPYKGTGTANPKVTKFMSEIMIDAESGRIDVIDKELTOPFEILFEGVY 361
Db 289 SSQSAGRDYEBYRFRAGTPTNAKSLKLVQFRUSENLRTDVCIKELQCLPTFAFWLEY 348
QY 362 IARAGWTEGKYAWSILLDRSQTRQLQVLISPELFI-----PVEDDMERQRLIE 411
Db 349 IYRVGWTDSRYVWAQLLDRPQQRLELVLLPVDNFCFIYSSSSSHPRKSSGWSRPLDK 408
QY 412 SVPDSTVEL-IYBETTDIMINHDIPHVPQSHSEIEPIFAPASECKTCFRHLHYKLTSL 470
Db 409 S---TTRPLQVIYTTETSSWVNVHVDLQ-EVELSEQEVTFMASE-ESGRHLHYLVISL 463
QY 471 KESKYRSRSGGLPAPSPDFKCP-----IKEEIAITSGEWEVLRHGSNIOVDEV 518
Db 464 SPNEV--SSGGVGA-TDHSLSMACIGSTLVARIVQKVTLTGGDWEVLGR--NVWYDRA 517
QY 519 RRLVVFEGTKDSPLEHLYVYVYVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKP 578
Db 518 RQLVYFMGLRETPELEKHLVYVSLAQPNQLRLTMPGYSFTVEFNDDCTLFLQTCYNISTL 577
QY 579 HCVSLYKL---SSPEDDPTCKTKFMAITLDSAGPLDYTP----- 616
Db 578 PSWELVRIAHDSNTANGNC-----SHGTPPTPIDALRLCSVGYLTEGGPSE 626
QY 617 -----PEIPESTTGTFTLYGMLYKPHDLQPKKYPTVLFYGGPOVQLVNNRFGVKYF 671
Db 627 NTQYNPSIHSPOISSGDVLYAMVFKPHFNLGVKPYPTLVNLYGSGPEVQTVTSNFKGMROL 686
QY 672 RLNTLASQYVVVVDNRGSGHGLKKEGAPKYKMGQIEIDDOVEGLQYLASRYDFIDL 731
Db 687 RMHMLASQYCVICVDSRGRHGVFEFESYRCRMGTVELSDQVEVLRLADQLGVIDMD 746
QY 732 RVGIHWSYGGVLSLMALMQRSDTRVAIAGAPVTLFIYDVTGYTERYMGHPQDQEGY 791
Db 747 RVAIHWYSYGGVLSLMLGVLPPEIPKSIAGAPVTSWEYDVTGYTERYMDLPDSNRSGYA 806
QY 792 LGSVMAQAKFPSEPNRLLLHGLDENVHFAHTSILLSELVRAGKYDYLQIYPOERHSI 851
Db 807 AGSVLNYIQKFPDENRLLIHLGLDENVHFAHTSILLSELVRANKYQIYQVYFNERHSL 866
QY 852 RVPESGEYELHLLHYLQEN 871
Db 867 RNLEASKHYETKLLSFLQNH 886

RESULT 9
Q9VC20 DROME PRELIMINARY; PRT; 1053 AA.
ID Q9VC20 DROME PRELIMINARY; PRT; 1053 AA.
AC Q9VC20; Q8IIH07;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG3744-PA, isoform A (CG3744-pc, isoform c) (LD33755p).
GN ORFNames=CG3744;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers J.-H., C. Blazer R.G., Champe M., Pfaffner B.D.,
RA Wan K.H., Doyle C., Baxter E.C., Holt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-U., Andrews-Pfannkuch C., Baldwin D.,
```


RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	RP	NUCLEOTIDE SEQUENCE.
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	RG	FlyBase;
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	RL	Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	RN	
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	[7]	NUCLEOTIDE SEQUENCE.
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	RC	STRAIN=Berkeley.
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	RA	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,	RA	Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Frise E.,
RA	Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	RA	George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA	Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,	RA	Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,	RA	Celniker S.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	RA	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RA	Kimmel B.E., Kodira K., Kraft C., Kravitz S., Kulp D., Lai Z.,	CC	-!- INTERACTION:
RA	Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,	CC	QVZRO:CG12016; NBExp=1; IntAct=EBI-105926, EBI-152819;
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	DR	EMBL; AE003749; AAF56357.2; -; Genomic_DNA.
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,	DR	EMBL; BT001499; AAN71254.1; -; mRNA.
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	DR	IntAct; Q9VC20; -;
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,	DR	MEROPS; S09.016; -;
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	DR	FlyBase; FBgn0039240; CG3744.
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	DR	FlyBase; FBgn0039240; CG3744.
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	DR	GO; GO:0016020; C:membrane; IEA.
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	DR	GO; GO:0003824; F:catalytic activity; IEA.
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,	DR	GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,	DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	DR	InterPro; IPR001375; Peptidase S9.
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,	DR	InterPro; IPR002469; Peptidase S9B.
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,	DR	InterPro; IPR003179; Ser. esters.
RT	"Finishing a whole-genome shotgun: release 3 of the Drosophila	DR	Pfam; PF00930; DPPIV.N.1.
RT	melanogaster euchromatic genome sequence."	DR	Pfam; PF00326; Peptidase S9; 1.
RL	Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).	DR	Pfam; PF00326; Peptidase S9; 1.
RL	Science 287:2185-2195(2000).	SQ	SEQUENCE 1053 AA; 118053 MW; C94AA663AB464577 CRC64;
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=22426065; PubMed=12537568;		
RA	Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,		
RA	Patel S., Adams M., Champe M., Dugan M., Lavery T., Muzny D.M., Nelson C.R.,		
RA	George R.A., Hoskins R., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,		
RA	Pacleb J.M., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,		
RA	Svirskas R., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,		
RA	Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,		
RT	"Finishing a whole-genome shotgun: release 3 of the Drosophila		
RT	melanogaster euchromatic genome sequence."		
RL	Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).		
RL	[3]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=22426070; PubMed=12537573;		
RA	Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,		
RA	Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,		
RA	Ashburner M., Celniker S.E.,		
RT	"The transposable elements of the Drosophila melanogaster euchromatin:		
RT	a genomic perspective."		
RL	Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).		
RL	[4]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=22426069; PubMed=12537572;		
RA	Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,		
RA	Gradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,		
RA	Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,		
RA	Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,		
RA	Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,		
RA	Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,		
RA	Lewis S.E.,		
RT	"Annotation of the Drosophila melanogaster euchromatic genome: a		
RT	systematic review."		
RL	Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).		
RL	[5]		
RP	NUCLEOTIDE SEQUENCE.		
RG	Berkeley Drosophila Genome Project;		
RA	Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,		
RA	Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,		
RA	Yu C., Rubin G.,		
RT	"Drosophila melanogaster release 4 sequence."		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		

QY 468 SILKESKYRSGGLPAP-----SDFKCPKEIEIATSGEWELGRHGSNTQV 515
DB 633 ASLLLSQ-----ANGQPDGSGVGAQPSFVDSLALQPRFLNKVAUTSGEWELAR---NLWV 685
QY 516 DEVRRLVYFEGTKQSPLEHLLVYVYNPGEVTRLTDRGYSHSC-----CISQH 564
DB 686 DRANKLVYFVGLRDTPEKHLVYVVSLEPSPHRIILTEPGSYLVLEFDDHFNDLFPFISQ 745
QY 565 CDFEISKYSN-QKNPHCVSLYKLSSPDDPTCKTFKFWATILDSAG-----PLPDYTP 616
DB 746 CKMLLVYCNIOQLPSCKVM-----RVNQTCSNGVNGVQIQLSVGLYHGGKPPQYCV- 798
QY 617 PEIFSEFTGTGLVGLYKPHDLQPKGYPTVLFIYGGPQVLVNNRFGVKYKRLNLT 676
DB 799 POIFSPQLPSGDIYVAVFKPHNFELGVKYPTVLNVYGGPEVQTVNNTFKGHLRMEML 858
QY 677 ASLGVVVVINDRSGCHRGKFGAFKVKMGQLEIDDOVEGLQYLASRYDFIDLORVGH 736
DB 859 AAGQYCVICIDSGSRHGRKFPESHIRGRMGQVELTQVDALRSLDQLGYIDMRVAIH 918
QY 737 GHSYGGYSLMALQORSDFRVAITAGAPVTLWTFYDGYTYRYMGHPDQNEOGYVLGSA 796
DB 919 GMSYGGYSLMGLVQYPIKFAITAGAPVTLWTFYDGYTYRYMGHPDQNEOGYVLGSA 978
QY 797 MQAEKPPSPNRLHLLHGFLENVHPHATSILLSFLVRACKPYDLOIYQERHSTRVPS 856
DB 979 EYVNPFPBEDKRLHILHGLIDENVHFCHTSLRSLANKANKPYEVHLFPFERHSLNLES 1038
QY 857 GEHYELHLLHYLQ 869
DB 1039 NKNYETKLLSLFQ 1051
RESULT 10
ID QVCL19 DROME PRELIMINARY; PRT; 1113 AA.
AC QVCL19;
AT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG3744-PB, isoform B.
GN Names=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phyllophoridae; Drosophilidae; Drosophila.
OX NCBI_taxids=727;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Li P.W., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Davenport L.B., Davies P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Golek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson K.A., Nixon K., Nuskern D.R., Paclebb J.M., M.G.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Purv V., Reese H.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Spradling A.C., Simpson M., Stappleton M., Smith T.,
RA Spier E., Spradling A.C., Stappleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhao L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Paclebb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirkas R., Tabor P.E., Wan K., Stappleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stappleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stappleton M., Paclebb J., Park S., Svirkas R., Smith E.,
Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003749; AAF56356.2; -; Genomic_DNA.
DR MEROPS; S09.016; -;
DR Ensembl; CG3744; Drosophila melanogaster.
DR FlyBase; FBgn0039240; CG3744.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser. esters.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.

```
SQ SEQUENCE 1113 AA; 123881 MW; 20857E3B212DF2E4 CRC64;

Query Match 34.0%; Score 1596; DB 2; Length 1113;
Best Local Similarity 37.9%; Pred. No. 5e-107;
Matches 369; Conservative 159; Mismatches 307; Indels 138; Gaps 27;

QY 2 AAAMEYEQI-----GVEIFETAD--CEENTIESQDRPKLEFFVY-----ERYWSOLK 46
DB 172 AASVITHLSSPTSGTPPHGLVDGDDCEDDEDDPVDNDGHIAAPTPNKNWAENK 231
QY 47 KLLADTKYHGVMAKAPDFMFKVRNPDGPHSDRIYIYLAWSGENRENTLFYSEIPKTI 106
DB 232 QVQVEIRKMCNLSWPTNVQF--RHLSDG--RARCFLGTTPQSWETILLFADINLTQ 287
QY 107 NRAAVLM-----LSWKPLLDL--FQAT 126
DB 288 SEEQQLLVQRLGIADEWSPPTWAGSPITSSGHQPAFLFNSLPRLPWSPLLOQPQSS 347
QY 127 LDYG-----MYGREBELLRKRIGTVGSIADYVHOGSGTFLFOAGSGIYHVKGQPG 180
DB 348 GSGGSSASPYAREYQLQERKLSWTGITSYELHKPSGKLVFPFCFNDLYQCLD---TG 404
QY 181 FTQOPLRPNLVETSCN--TRMDPKLCPADPDWIAFHNSNDIWNISVTRREERLTYVHNE 239
DB 405 YNSGLLPFTQLRT--CPQWALDPOICPQNSDMIAYISDCDLFVTHLSGHEKRLTYTSTG 463
QY 240 LANMEEDARAGVATFVLQEEFDRIYGYWCKAETTSYGKILIRILYEENDESEVLIH 299
DB 464 RHSVYDDALSAGVPSYVMOEFSRYQGFWMQPH-----SNDGIYRIYVEEVESESVYT 518
QY 300 VTSPLMLETADSFPRPKGTANPKVTFKWSIMIDAEGRIIDVIDKELIQPREIL--PE 357
DB 519 FPSSTAMHGRVDEYRPRPTGSPNASKLKVQFLNEALQVSEIAKDL--PYSLLA VFS 576
QY 358 GVEIARAGWTPGKAWSTLDRSOTRIQIVLISPELFIPEVDDVMERQRLIESVPD-- 415
DB 577 WLEYIVRVGWTPOAKYVVOGLDRKQOORLDVILILPDLNF--CESYSSQVSTPTDSIGDHS 634
QY 416 -----SVTPL--IIYEETTDIWINIHDFHVPFQSHHEEIEFIPASECKTGFRHLYKIT 467
DB 635 WRSLSYSTIIPLOVITERSDSWINVHMLH--FLDITSVTFLWASE--ETGFRHLYVT 692
QY 468 SILKESYKRSSGGLPAP-----SDFKCPKEEIAITSGEWEVLGRGSGNIQV 515
DB 693 ASLLLSQ-----ANGQPDGSGVGAQPSFVLSALQPRILNKVALTSGEWEVLAR---NLW 745
QY 516 DEVRALVYFGTKDPSLEHLLHYVSVYVNGEVTRLTDRGYSHSC-----CLSQH 564
DB 746 DKANKLVYFVGLRDTPLEKHYVVSLSRPERHRLLTLEPGSYLVBEFDDHFNDFPFISQ 805
QY 565 CDFIFISKYSN--QKNPHCVSLYKLSSEDDPTCKTEFWATILDSAG-----PLPDYTP 616
DB 806 CKMLLVYCNITQLPSCKNV-----RVNQTCNSGVNGIQISLVGYLHEGKPEPQYC- 858
QY 617 PEIYSPSTTGFTLYGMLYKPHDLQPKKYPTVLFIYGGPQVQLVNNRFKGVYFRNLTL 676
DB 859 PQIFSPQLSGDIYVAMVFPNHPNFGVYPTVNLVYGGPEVQVNTNTFKGKHQLRHML 918
QY 677 ASLGVVVVINDGSGCHRGKLFEGAFKYMGOIIDDQVEGQYLAISYDFIDLDVIGIH 736
DB 919 AAQGYVICIDSGSRHRGKRFESHIRGRMGQVELDQVDAURLSLSQGLYIDMDRVAIH 978
QY 737 GWSYGGYLSLMALMORSDIFRVAIAGAPVTLMIFFDTGYTERYMGHPDQNGEQYLGSA 796
DB 979 GWSYGGYLSLMGLVQYFKI FKVAIAGAPVNTWEYDTGYTERYMDMPQNNQAGYSAGSVL 1038
QY 797 MQAKFPSPENRLLHGLFDENVHFAHTSILLSFLVRAGKYDIOIYQPBHRHSIRVPS 856
DB 1039 EYVNSFPEDKRLLLIHLIDENVHFCHTSRLISALNANKPKYEVHLFPEERHSRLNLES 1098
QY 857 GEHYELHLLHYLO 869
DB 1099 NKNYETKLLSFLQ 1111
```

RESULT 11

```
Q5TTK8 ANOGA
ID Q5TTK8 ANOGA PRELIMINARY; PRT; 740 AA.
AC Q5TTK8;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE ENSANGP00000026132 (Fragment).
GN ORFNames=ENSANG00000014037;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG "The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008879; EAL40717.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase S9B.
DR InterPro; IPR000379; Ser esters.
DR Pfam; PF00930; DPPIV N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON TER 1
SQ SEQUENCE 740 AA; 84151 MW; D59F131BBA095CCB CRC64;

Query Match 33.3%; Score 1567; DB 2; Length 740;
Best Local Similarity 42.3%; Pred. No. 3.6e-105;
Matches 330; Conservative 129; Mismatches 241; Indels 80; Gaps 15;

QY 133 SREBELRERKRIGTVGSIADYVHOGSGTFLFOAGSGIYHVKGQPGQFTQOPLRPNLVE 192
DB 1 SREVOLMERKRUSIGWITSYELHKASGKIIVFPACNTLYQCLD---TGYEENPLFPSEL 57
QY 193 TSCPNIRMDPKLCPADPDWIAFHNSNDIWNISVTRREERLTYVHNELANMEEDARSAGV 252
DB 58 ILQRAAALDPOICPQNSDLVAFVCGDIWVHTSHGSHSERLTAYHDGRRSFAEDPLTAGV 117
QY 253 ATFVLQBEFDRYGYWVWCPKAETTPSGGK-----ILRILYEENDESEVEIILHVTSP 303
DB 118 PSYVWQEFSTRYQGFWMQPE-----SHGKFDMDVILDEVYRIVYEVEDSVLTFPSS 172
QY 304 MLETRRADSPRYKPTGTANPKVTFKMSIEMDAEGRIIDVIDKELIQPEILFEGVEYIA 363
DB 173 QSAGRDYEEYRFPAGTPNAKSKLVQFRLSENRLITDVCIKELQCPLTFAPPWLEIV 232
QY 364 RAGWTPGKYAWSILLDRSOTRIQIVLISPELFIPEVDDVMERQRLIESVP-----DS 416
DB 233 RVGTPDSRYVWAAQLDRPQORLELVLLPVDNFCETYS-----SSRPPAGARRKST 284
QY 417 VTPLE-IIYEETTDIWINIHDFHVPFQSHHEEIEFIPASECKTGFRHLYKITSLKESKY 475
DB 285 TRPLQVIYTTETSSSWNVNVDVLQ-FVELSSEQVETFLWASE--ESGFRHLYLTVS----- 335
QY 476 KSSSGGLPAPDFKCP-----IKEEIAITSGEWEVLGRHSGNIQVDEVRRLVY 523
DB 336 KVSSGGVGGA--TDHSLPSMACIGSTLVARIQVKVTLTGDEWELGR---NWYDRAQLVY 391
```



```
DR GO: 0003824; F: catalytic activity; IEA.
DR GO: 0008236; F: serine-type peptidase activity; IEA.
DR GO: 0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro: IPR001375; Peptidase_S9.
DR InterPro: IPR002469; Peptidase_S9B.
DR InterPro: IPR000379; Ser esters.
DR Pfam: PF00930; DPPIV_N; I.
DR Pfam: PF00326; Peptidase_S9; 1.
FT NON_TER 1
FT NON_TER 557
SQ SEQUENCE 557 AA; 62685 MW; 08AAAA944ACD0808 CRC64;

Query Match 24.3%; Score 1142.5; DB 2; Length 557;
Best Local Similarity 43.3%; Pred. No. 2.2e-74;
Matches 253; Conservative 84; Mismatches 198; Indels 49; Gaps 12;

QY 284 RILYEENDESEVEIHTVTSPLMSTRADSFYPKTKGTANPKVTPKMSIMIDABGRIDV 343
Db 1 RIVYEEDSDVSLYTFPPSQSAGRDYBYRFRAGTPNNAKSLKLVQFRLENLRITDV 60
QY 344 IDEKLIQFFELFEGVEIYARAGWTEPGKYAWSILLDRSQTRQLQVLISPELIPVEDDV 403
Db 61 CIKELQCLPLFAFPWLEYIVRGWTPDSRYVWAQLLDRPQORLELVLPVDFCSTA--A 118
QY 404 MERQRLIESVPDSVTPLIIVYEETDIIWINIHDFHVFPQSHHEEIEFIPASECTGPRHL 463
Db 119 RRRPTAGRPAGGAMGPGIYTTSSWVNVHVDLQ--FVELSDEY-----TTNSHP 168
QY 464 YKITSILKESKY--KRSSGGLPAPDFKCP-----IKKEIATISGEVNLGRH 509
Db 169 TAPKQKQREGAHTGVSSGAGA--TDHSLPSPMACIGSTLVARIVQKVLITGQDWEVLGR- 226
QY 510 GSNLOVDEVRLLVFECKDPSLEHLYVSYVNPCEVTRLTDRGYSHSCCISOHCDPF- 568
Db 227 --NVYDRVRLVQVYFVGLRETPLEKLYVYVLAQNPQLRLTTPGYSFTV---BFNDLVR 281
QY 569 ISKYSNQNPKHCVSLYKLSPEDDPTCKTKFEWATILDSAGLPDYTP-----PEIFS 621
Db 282 IAHDSNTANGCS-----HGTPPTPT-----PIDALRCLSVGLTEGGSENTQNPISHS 333
QY 622 FESTTGTFLYMLYKPHDLQPKKYPVTLFIYGGPOVLVNNRFGVKYFRLNTLASLGY 681
Db 334 PQISGGDLVYANVFKPHFNLGKYPVTLVNVYGGVEQVQVNTFKGMQLRWHMLASQGY 393
QY 682 VVVVIDNRGSGHRLKPEGAFKYMGOIEIDDOVEGLQVLAIRYDFIDLDLVGHGWSYG 741
Db 394 CVICVDSRGRHGRVFEFESYIRRRMGTVLSQVEVLRILADQLGYIDMDRAVIAHGWSYG 453
QY 742 GYLSMALMQSDIFRVAIAGAPVTIWIYDTGTERTYRNGHPDQNEQGYILGSVAMQAEK 801
Db 454 GYLSLMLGLVQYPEIFKVSIAAGAPVTSWEYDGTERTYRMDLPDSNRSGYAAAGVNLVYIQ 513
QY 802 FPFSEPNRLLLHGLDENVHFAHTSILLSLFLVRAGKPYDLQYYP 845
Db 514 FPDENRLLIIGHLDENVHFHTSQLVSLRVLANKPYQLQYYP 557

RESULT 14
Q54U01.DICDI PRELIMINARY; PRT; 803 AA.
AC Q54U01;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD0205566;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,

RA Sugang R., Bertman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tungal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kethornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsey R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Louissegh H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Williams R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF10100074; EAL66689.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 803 AA; 93187 MW; 33846B370C039FAE CRC64;

Query Match 23.9%; Score 1124.5; DB 2; Length 803;
Best Local Similarity 33.0%; Pred. No. 7.9e-73;
Matches 284; Conservative 142; Mismatches 299; Indels 136; Gaps 29;

QY 61 AKAHDF-----MFVKNDPDGPHSDRIYILAMSGENRENTLIFYEIPKTNRAVL 112
Db 30 SSAPVSFEFNDKSNHLYFLSNVKDKQNKIHIYIDNMSDSKEIKFLFNVVDKDE- 84
QY 113 MLSWKPLDLFOATLDYGMYSREELRERKIGTIGTASYDYHQSGTFLFQAGSIYH 172
Db 85 -----LSIEDQLQRERMTAANGITQFTFDQKHQFIAPINNKINK 125
QY 173 VKDGGPOGFTQOPLRPNLVETSCPNRMPDKLPADPDWIAFIHSDNWSINIVREERR 232
Db 126 IDI--KESITKPIKTIIVGETYNHQSADGKI-----VSLKDKIDWIDTISNAMYR 176
QY 233 LTVVHNLNANNEEDARSAGVATFVLOBEFDYSGYWCPCAKETTPSGGK-ILRLIYEND 291
Db 177 ITFSNDE---KHKFRVAGDIGIYAEESRVTGYWSPVIGTCVTKGKPMYTCYLBED 233
QY 292 ESEVEIHTVSPMLETR-RADSPRYPKTGTAN-----PKVTFKMSIEMDAE 337
Db 234 ETNWDYHI--PTSDLRGKTKYKYPKLAGKNSICKVCLVSFVLTRTTFQDSK----- 285
QY 338 GRIIDVIDKELIQFELFEGVEIYARAGWTEPGKYAWSILLDRSQTRQLQVLISPELFI 397
Db 286 ---IEIVKSELFD-LKTFPPWAEVITRAGWTPNGHSIYLLQLLDRKQOHLVWPLHVF- 340
QY 398 PVEDDWMERQRLIESVPDSVTPLIIVYEETDIIWINIHDFHVFPQSHHEEIEFIPASECK 457
Db 341 -AEDYSSSSSSSVKSIKPK--LPVLI-EETTSVINIEFSFQ-FLKSINQL--IWSNE-Q 392
QY 458 TGFPHLYKITSILKESKYKSSGGLPAPSDFKPIKEEITAITSGEWEVLRHGSIQVDE 517
Db 393 SGYRHLV----LIKWDKNFTNIQSTPTILS-TCNDNDN---DNNNMV---SSDDIHIDE 441
QY 518 VRLVYPEGTKDSPLEHLYVWVSNVP-GEVTRLTDRGYSH-SCCISQHCDFFIKYSNQ 575
Db 442 KRKLVIYFTGTDCTCLEOHLIVTRFDKPNSEIKKLSHANFSHSISISSNFKFITYSNI 501
QY 576 KN-----PHCVSLYKLSPEDDPTCKTKFEWATILDSAGLPD 613
Db 502 STISKTEVFDLIYNDNDNDNDNDIYPIVKSFFINDDDDDDDKKKININI----- 552
QY 614 YTPPEIFSFEFTGTFLYGLMKYKPHDLQPKKYPVTLVFIYGGPOVLVNNRFGVKYFRL 673
Db 553 ---PKIFNFKNSKGVTIYGVYTLPLPSDYKDKKYPVTVVYVGGPHVQIVRNOYNIKQHYT 609
```

QY 674 NTLASIGVYVVVVIDNRGSGHRLKFGAFKYMKGQIEIDDQVEGLQYLASRYDF--IDLQ 731
Db 610 N-----FGFIQVIMDNVGSANRGLFESHIREFKMGQVEIGDQVEGINYLGN-DIVSIDVN 664

QY 732 RVGIHGWSVGGYLSLMLMQRSDFRVAIAGAPVTWIFDYTGTRYVMGHPDQNEQGY 791
Db 665 RTAISGWSVGGYNSLMAISQRPDVFVKIACGAPVSDWRLYNTGTRYVMNVPQNDIDGYK 724

QY 792 LGSVAMQAEKPESEENRLLHGLDENVHFHTSILLFSLVRAGKPYDLQIYPOERHSI 851
Db 725 LGDTTHYS--FTEENRLLIHGLQDENVHFSNTIYIIDHLTKTKPYILKTLPLNERHGV 782

QY 852 RVPESGEHYELHLHYLOENL 872
Db 783 RNTDNRIYIGLFWINHLKLN 803

RESULT 15
O75868 HUMAN
ID O75868 HUMAN PRELIMINARY; PRT; 432 AA.
AC O75868;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE R33083_1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Attix C., Andrise T., Trankheim M., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Quan G., Kronmiller B.,
RA Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrasco A.V.,
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC005783; AAC62840.1; -, Genomic DNA.
DR Ensembl; ENSG00000142002; Homo sapiens.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DPPIV_N; 1.
FT NON_TER 432
SQ SEQUENCE 432 AA; 48595 MM; 64E2B85BE0523A7E CRC64;

Query Match 22.3%; Score 1049; DB 2; Length 432;
Best Local Similarity 55.8%; Pred. No. 1e-67; Indels 8; Gaps 2;
Matches 189; Conservative 63; Mismatches 79;
QY 35 FYVERYSQKLLADTRKYHYGMMAKAPHDFMFKRNDPDGPHSDRIYYLAMSNGRE 94
Db 95 FQVKHSDGLRSIIHGRKYSGLIYNKAPHDFQVQKTDSGPHSHRLYYLGMPSGRE 154
QY 95 NTLFYSEIPKTIINRAVLMLSWKPLDLDFQATLDYGMYSREELRLRKRKIGTVGIASYD 154
Db 155 NSLLYSBIPKKVRKEALLLSWKQMLDHFQATPHGVYSREELRLRKRKLGVEGITSYD 214
QY 155 YHOGSGTFLFQAGSGIYHVKGQPGF-----TQQLRPNLVETSCPNIRMDPKLCPA 207
Db 215 FHSESGFLFQASNSLPHCRDGGKNGFMVSPGCVSPMKPLEIKTCGSRMDPKICPA 274
QY 208 DPDIATFHSNDIWSINVTREERRLLTYVHNELANNEEDARSAGVATVLOEEFDYSGY 267
Db 275 DPAPFSFINSDLVANIEGERRLLTFCHQGLSNVLDPPKSAGVATFVIOEEFDRTGY 334
QY 268 WPCPKAETFGGKILRLIYEENDESEVEIHHVTSMLFTRADSFRYPKGTGTANPKVTF 327
Db 335 WACPTA-SWEEGLKTLRLIYEVEDESEVEIHHVTPSPALERKTDSDYRYPRTGSKNPKIAL 393

Search completed: April 14, 2006, 12:50:08
Job time : 246 secs

QY 328 KMSEIMDAEGRIIDVIDKELIQPFELPEGVEYIARAG 366
Db 394 KLAEPQDTSQKIVSTQEKELVQPFSSLPFKVEYIARAG 432

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 12:50:22 ; Search time 49 Seconds
(without alignments)
1488.162 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMETEQLGVEIFETADC.....HLLHYLQNLGSLRIALKVI 882

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/1aa/5 COMB.pep: *
2: /cgn2_6/prodata/1/1aa/6 COMB.pep: *
3: /cgn2_6/prodata/1/1aa/H COMB.pep: *
4: /cgn2_6/prodata/1/1aa/PCTUS COMB.pep: *
5: /cgn2_6/prodata/1/1aa/RE COMB.pep: *
6: /cgn2_6/prodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	100.0	882	2	US-09-976-674-1
2	4700	100.0	882	2	US-10-070-464-1
3	3504	74.6	658	2	US-09-976-674-19
4	3504	74.6	661	2	US-09-976-674-11
5	3504	74.6	690	2	US-09-976-674-7
6	3236	68.9	613	2	US-09-976-674-21
7	2870	61.1	863	2	US-09-976-674-3
8	2870	61.1	892	2	US-09-976-674-23
9	2870	61.1	892	2	US-09-976-674-27
10	2820.5	60.0	879	2	US-09-976-674-33
11	2820.5	60.0	879	2	US-09-976-674-35
12	2422	51.5	465	2	US-10-070-464-5
13	2406	51.2	832	2	US-09-976-674-29
14	2406	51.2	832	2	US-09-976-674-31
15	2356.5	50.1	819	2	US-09-976-674-37
16	2356.5	50.1	819	2	US-09-976-674-39
17	1836.5	39.1	360	2	US-10-070-464-7
18	1808	38.5	358	2	US-09-976-674-13
19	1645.5	35.0	310	2	US-09-794-236-4
20	1645.5	35.0	310	2	US-10-070-464-3
21	1391	29.6	518	2	US-09-976-674-25
22	1278	27.2	241	2	US-09-976-674-9
23	1007.5	21.4	194	2	US-09-976-674-17
24	719	15.3	981	2	US-09-902-540-16812
25	616	13.1	710	2	US-09-518-550-28
26	582.5	12.4	771	2	US-09-462-284-2
27	582.5	12.4	771	2	US-09-079-592-2

28	572.5	12.2	723	2	US-09-518-550-29	Sequence 29, Appl
29	572	12.2	732	2	US-09-518-550-30	Sequence 30, Appl
30	569.5	12.1	676	2	US-09-518-550-42	Sequence 42, Appl
31	529	11.3	766	2	US-10-002-593-6	Sequence 6, Appl
32	529	11.3	766	2	US-09-949-016-6146	Sequence 6146, Ap
33	529	11.3	766	2	US-10-423-714-6	Sequence 6, Appl
34	525	11.2	755	4	PCT-US93-07923-3	Sequence 3, Appl
35	525	11.2	759	4	PCT-US93-07923-2	Sequence 2, Appl
36	525	11.2	766	1	US-08-230-491A-3	Sequence 3, Appl
37	525	11.2	766	1	US-08-619-280A-3	Sequence 3, Appl
38	525	11.2	766	1	US-08-940-391-3	Sequence 1, Appl
39	525	11.2	766	2	US-09-794-236-1	Sequence 3, Appl
40	525	11.2	766	2	US-09-265-606-3	Sequence 10450, A
41	525	11.2	775	2	US-09-949-016-10450	Sequence 27, Appl
42	519	11.0	766	2	US-09-518-550-27	Sequence 11421, A
43	488.5	10.4	737	2	US-09-902-540-11421	Sequence 3, Appl
44	484.5	10.3	818	2	US-09-462-845-3	Sequence 3, Appl
45	484.5	10.3	818	2	US-10-402-312-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Query Match	100.0%;	Score	4700;	DB 2;	Length	882;	
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	882;	Conservative	0;	Mismatches	0;	Gaps	0;
Qy	1	MAAMETEQLGVEIFETADCEENIESQDRPKLEPFVVERYSWSQLKLADTRKYHGYYMM	60				
Db	1	MAAMETEQLGVEIFETADCEENIESQDRPKLEPFVVERYSWSQLKLADTRKYHGYYMM	60				
Qy	61	AKAPHDMFVKRNDPGPHSDRIYYLAMS GENRENTLFYSEIPKTIINRAAIVLMSWKPLL	120				
Db	61	AKAPHDMFVKRNDPGPHSDRIYYLAMS GENRENTLFYSEIPKTIINRAAIVLMSWKPLL	120				
Qy	121	DLFQATLDYGMYSREELLRRKRIGTVGIASVDYHQSGGTFLFOAGSGIYHVXDGQPGQ	180				
Db	121	DLFQATLDYGMYSREELLRRKRIGTVGIASVDYHQSGGTFLFOAGSGIYHVXDGQPGQ	180				
Qy	181	FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWI SNITVTRERRLTYVHNEL	240				
Db	181	FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWI SNITVTRERRLTYVHNEL	240				
Qy	241	ANMEEDARSAGVATFVLQEEFDYSGYWCPCAETTPSGGKILRLIYENDESEVEIHHV	300				
Db	241	ANMEEDARSAGVATFVLQEEFDYSGYWCPCAETTPSGGKILRLIYENDESEVEIHHV	300				
Qy	301	TSPMLETRRADSPRYPKTGTANPKVTFKQSEIMDAEGRIIDVIDKELIQPFELFEGVE	360				
Db	301	TSPMLETRRADSPRYPKTGTANPKVTFKQSEIMDAEGRIIDVIDKELIQPFELFEGVE	360				

QY	361	YIARAGWTPEGKYAWSILLDRSOTRLQIVLISPELFI	PVEDDDVMERQRLIESVPDSVTPL	420
Db	361	YIARAGWTPEGKYAWSILLDRSOTRLQIVLISPELFI	PVEDDDVMERQRLIESVPDSVTPL	420
QY	421	IYVEETTDIWINIHDIHVFPQSHREEIEFI	FASECKTGFRHLYKITSLKESKYKRSG	480
Db	421	IYVEETTDIWINIHDIHVFPQSHREEIEFI	FASECKTGFRHLYKITSLKESKYKRSG	480
QY	481	GLPAPSDPKCPDKBEIATISGEWEVLGRHGSNI	OVDEVRLVYFEGTKDSPLEHLYVVS	540
Db	481	GLPAPSDPKCPDKBEIATISGEWEVLGRHGSNI	OVDEVRLVYFEGTKDSPLEHLYVVS	540
QY	541	YINPGEVTRLDRGSHSCCISQCHDFI	SKYSNOKNPHCVSLYKLSSEDDPTCKTEF	600
Db	541	YINPGEVTRLDRGSHSCCISQCHDFI	SKYSNOKNPHCVSLYKLSSEDDPTCKTEF	600
QY	601	WATILDSAGPLPDYTPPEIFSFESTT	GFTLYGMLYKPHDLQPKKYPTVLFIYGGPQVL	660
Db	601	WATILDSAGPLPDYTPPEIFSFESTT	GFTLYGMLYKPHDLQPKKYPTVLFIYGGPQVL	660
QY	661	VNNRFGVKYFRLNTLASLGYVVVVV	IDNRGSHRGLKFEFAGFYKMGQIEIDDOVEGLQY	720
Db	661	VNNRFGVKYFRLNTLASLGYVVVVV	IDNRGSHRGLKFEFAGFYKMGQIEIDDOVEGLQY	720
QY	781	GHPDQNEQGYILGVSAMQAEKFP	PEPNRLLHLLHGYLQENLGSRIAALKVI	882
Db	781	GHPDQNEQGYILGVSAMQAEKFP	PEPNRLLHLLHGYLQENLGSRIAALKVI	882
QY	841	LOIYPQERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI	882	
Db	841	LOIYPQERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI	882	
RESULT 2				
US-10-070-464-1				
; Sequence 1, Application US/10070464				
; Patent No. 6881564				
; GENERAL INFORMATION:				
; APPLICANT: ABBOTT, Catherine Anne				
; APPLICANT: GORRELL, Mark Douglas				
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES				
; FILE REFERENCE: GH-007				
; CURRENT APPLICATION NUMBER: US/10/070,464				
; CURRENT FILING DATE: 2002-03-07				
; PRIOR APPLICATION NUMBER: PCT/AU00/01085				
; PRIOR FILING DATE: 2000-09-11				
; PRIOR APPLICATION NUMBER: AU PQ5709				
; PRIOR FILING DATE: 2000-02-18				
; PRIOR APPLICATION NUMBER: AU PQ2762				
; PRIOR FILING DATE: 1999-09-10				
; NUMBER OF SEQ ID NOS: 8				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 1				
; LENGTH: 882				
; TYPE: PRT				
; ORGANISM: Homo Sapiens				
US-10-070-464-1				
Query March				
Best Local Similarity 100.0%; Score 4700; DB 2; Length 882;				
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MAAAMETEQIGVIFETADCEENIESODRPKLEPFVVERYSWSOLKLLADTRKYHYVM	60	
Db	1	MAAAMETEQIGVIFETADCEENIESODRPKLEPFVVERYSWSOLKLLADTRKYHYVM	60	
QY	61	AKAPHDFWVKRNDPDGPHSDRIYYI	LAMSGENRENTLFYSEIPKTI	NRAAVMLSWKPLL 120

LENGTH: 658
TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-674-19

Query Match 74.6%; Score 3504; DB 2; Length 658;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKHYGYM 60
Db 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKHYGYM 60

Qy 61 AKAPHDFMVKNDPDGPHSDRIYYIYVAMSGENRENTLYFSEIPKTIINRAAVLMLSWKPLL 120
Db 61 AKAPHDFMVKNDPDGPHSDRIYYIYVAMSGENRENTLYFSEIPKTIINRAAVLMLSWKPLL 120

Qy 121 DLFOATLDYGMYSREBEELRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPGQ 180
Db 121 DLFOATLDYGMYSREBEELRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPGQ 180

Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLLTYVHNL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLLTYVHNL 240

Qy 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEI 300
Db 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEI 300

Qy 301 TSPMLETRADSFYRPTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRADSFYRPTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360

Qy 361 YIARAGWTEGKYAWSILLDRSQRLQIVLISPELFIPEVDDVNERQLIESVPDSVTPL 420
Db 361 YIARAGWTEGKYAWSILLDRSQRLQIVLISPELFIPEVDDVNERQLIESVPDSVTPL 420

Qy 421 IYEEETDWINIHDFVFPQSHEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480
Db 421 IYEEETDWINIHDFVFPQSHEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480

Qy 481 GLPAPSDFKPCIKETAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDFKPCIKETAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540

Qy 541 YNPNGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
Db 541 YNPNGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600

Qy 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
Db 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655

RESULT 4
US-09-976-674-11
Sequence 11, Application US/09976674
Patent No. 6844180
GENERAL INFORMATION:
APPLICANT: Qi, Steve
APPLICANT: Akinsanya, Karen
APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976, 674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 661

TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-674-11

Query Match 74.6%; Score 3504; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKHYGYM 60
Db 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKHYGYM 60

Qy 61 AKAPHDFMVKNDPDGPHSDRIYYIYVAMSGENRENTLYFSEIPKTIINRAAVLMLSWKPLL 120
Db 61 AKAPHDFMVKNDPDGPHSDRIYYIYVAMSGENRENTLYFSEIPKTIINRAAVLMLSWKPLL 120

Qy 121 DLFOATLDYGMYSREBEELRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPGQ 180
Db 121 DLFOATLDYGMYSREBEELRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPGQ 180

Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLLTYVHNL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLLTYVHNL 240

Qy 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEI 300
Db 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEI 300

Qy 301 TSPMLETRADSFYRPTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRADSFYRPTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360

Qy 361 YIARAGWTEGKYAWSILLDRSQRLQIVLISPELFIPEVDDVNERQLIESVPDSVTPL 420
Db 361 YIARAGWTEGKYAWSILLDRSQRLQIVLISPELFIPEVDDVNERQLIESVPDSVTPL 420

Qy 421 IYEEETDWINIHDFVFPQSHEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480
Db 421 IYEEETDWINIHDFVFPQSHEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480

Qy 481 GLPAPSDFKPCIKETAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDFKPCIKETAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540

Qy 541 YNPNGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
Db 541 YNPNGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600

Qy 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
Db 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655

RESULT 5
US-09-976-674-7
Sequence 7, Application US/09976674
Patent No. 6844180
GENERAL INFORMATION:
APPLICANT: Qi, Steve
APPLICANT: Akinsanya, Karen
APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976, 674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 690
TYPE: PRT

```
; ORGANISM: Homo sapiens
US-09-976-674-7

Query Match
Best Local Similarity 74.6%; Score 3504; DB 2; Length 690;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETQLGVETAFETADCEENIESQDRPKLPFFVVERYSWSQLKLLADTRKHYGYMM 60
DB 1 MAAAMETQLGVETAFETADCEENIESQDRPKLPFFVVERYSWSQLKLLADTRKHYGYMM 60
QY 61 AKAPHDFMFKVXKNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNINRAAVLMLSWKPLL 120
DB 61 AKAPHDFMFKVXKNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNINRAAVLMLSWKPLL 120
QY 121 DLFOATLDYGYMSREEELLRRERKIGTVGIASDYDYGSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGYMSREEELLRRERKIGTVGIASDYDYGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOQLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREERLTYVHNEL 240
DB 181 FTQOQLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREERLTYVHNEL 240
QY 241 ANMEEDASAGVATFVLQEEFDRYSGYWCPCKAETTPSGGKILIRILYEENDESEVEIHHV 300
DB 241 ANMEEDASAGVATFVLQEEFDRYSGYWCPCKAETTPSGGKILIRILYEENDESEVEIHHV 300
QY 301 TSPMLETRRADSFYRYPKGTGTANPKVTFMSEIMIDAGRIIDVIDKELIQPFILFEGVE 360
DB 301 TSPMLETRRADSFYRYPKGTGTANPKVTFMSEIMIDAGRIIDVIDKELIQPFILFEGVE 360
QY 361 YIARAGWTPGEGKYAWSILLDRSOTRLQIVLISPELTFPVEDDVMERQRIIESVPDSVTPL 420
DB 361 YIARAGWTPGEGKYAWSILLDRSOTRLQIVLISPELTFPVEDDVMERQRIIESVPDSVTPL 420
QY 421 IYEEETDIIWINIHDIHFVFPQSHEEEIEFIPASECTGPRHLKYKITSILKESKYRSGG 480
DB 421 IYEEETDIIWINIHDIHFVFPQSHEEEIEFIPASECTGPRHLKYKITSILKESKYRSGG 480
QY 481 GLPAPSDFKCPKEBIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
DB 481 GLPAPSDFKCPKEBIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
QY 541 YNPGGEVRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
DB 541 YNPGGEVRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPKKYPVLFIYGG 655
DB 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPKKYPVLFIYGG 655

RESULT 6
US-09-976-674-21
; Sequence 21, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 21
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Homo sapiens

; ORGANISM: Homo sapiens
US-09-976-674-21

Query Match
Best Local Similarity 68.9%; Score 3236; DB 2; Length 613;
Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETQLGVETAFETADCEENIESQDRPKLPFFVVERYSWSQLKLLADTRKHYGYMM 60
DB 1 MAAAMETQLGVETAFETADCEENIESQDRPKLPFFVVERYSWSQLKLLADTRKHYGYMM 60
QY 61 AKAPHDFMFKVXKNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNINRAAVLMLSWKPLL 120
DB 61 AKAPHDFMFKVXKNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNINRAAVLMLSWKPLL 120
QY 121 DLFOATLDYGYMSREEELLRRERKIGTVGIASDYDYGSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGYMSREEELLRRERKIGTVGIASDYDYGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOQLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREERLTYVHNEL 240
DB 181 FTQOQLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREERLTYVHNEL 240
QY 241 ANMEEDASAGVATFVLQEEFDRYSGYWCPCKAETTPSGGKILIRILYEENDESEVEIHHV 300
DB 241 ANMEEDASAGVATFVLQEEFDRYSGYWCPCKAETTPSGGKILIRILYEENDESEVEIHHV 300
QY 301 TSPMLETRRADSFYRYPKGTGTANPKVTFMSEIMIDAGRIIDVIDKELIQPFILFEGVE 360
DB 301 TSPMLETRRADSFYRYPKGTGTANPKVTFMSEIMIDAGRIIDVIDKELIQPFILFEGVE 360
QY 361 YIARAGWTPGEGKYAWSILLDRSOTRLQIVLISPELTFPVEDDVMERQRIIESVPDSVTPL 420
DB 361 YIARAGWTPGEGKYAWSILLDRSOTRLQIVLISPELTFPVEDDVMERQRIIESVPDSVTPL 420
QY 421 IYEEETDIIWINIHDIHFVFPQSHEEEIEFIPASECTGPRHLKYKITSILKESKYRSGG 480
DB 421 IYEEETDIIWINIHDIHFVFPQSHEEEIEFIPASECTGPRHLKYKITSILKESKYRSGG 480
QY 481 GLPAPSDFKCPKEBIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
DB 481 GLPAPSDFKCPKEBIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
QY 541 YNPGGEVRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
DB 541 YNPGGEVRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDS 607
DB 601 WATILDS 607

RESULT 7
US-09-976-674-3
; Sequence 3, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-976-674-3
```



```
RESULT 9
US-09-976-674-27
; Sequence 27, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-27

Query Match      61.1%; Score 2870; DB 2; Length 892;
Best Local Similarity 61.5%; Pred. No. 1.2e-283;
Matches 517; Conservative 134; Mismatches 187; Indels 2; Gaps 2;

QY 35 FVVERYSQKLLADTRKYHGYMMAKAPHDEMFKRNDPDGPHSDRIYYLAMSGENRE 94
DB 53 FQVKHSDGLRSIIHGSRYKSLGVNKAHPDFQVQKTDSESPHSHRLYYLGMYPGSR 112
QY 95 NTLFYSEIPKTNINRAAVALMSKPLDLFOATLDYMGYSREBELLRKRIGTVGIASYD 154
DB 113 NSLLYSEIPKPKYKREALLLSWKQMLDHFQATPHHGYISREBELLRKRIGTVGITSYD 172
QY 155 YHQSGLTFLQAGSGIVHVDGPGQGTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAF 214
DB 173 FHSSEGLFLFOANSLSLHCRDGGKNGFPMVSPMKPLRIKTCQSGPRMDPKICPADPAFFSF 232
QY 215 IHSNDIWSNTVREERLTYVHNELANMEEDARSAGVATFVLQEBFDYSGYWCCKAE 274
DB 233 INNSDLWANTETGEERLTFCHQGLSNVLDPPKSGAGVATFVIOEBFDRFTGYWCPTAS 292
QY 275 TTPSGG-KILRLIYEENDESEVEIIVHTVSPMLTTRADSPRYPKTGTANPKVTFKMSIM 333
DB 293 WEGSEGLKTLRLIYEENDESEVEIIVHVPSPALBERKTDVSYRPTGSKNPKIALKLBQ 352
QY 334 IDAEGRIIDVDKELIOPPEILFEGVEYIARAGWTPGKYAMSILLDRSOTRLQIVLISP 393
DB 353 TDSQKIVSTQEKLVQPFSSLPFKVEYIARAGWTRDKYAWAMFLDRPQWLQVLLPP 412
QY 394 ELFIPEVDVMERQRLIESVPDSTPLIYYEETDWINHIDIFHVFPQSH-EEETEFIF 452
DB 413 ALFIPSTENESORLASARAVPRNVQPVVYEEVTNWINVHIDIFYPFPQSEGEDELCLFLR 472
QY 453 ASECKTGRPHLYKITSLIKESKYRSSGGLPAPSDPKPIKEEIAITSGEWELGRHGSN 512
DB 473 ANECKTGCHLYKYTAVLKSQGYDWSPEFPFGDEDFKCPIKEEIAITSGEWELABHGS- 532
QY 513 IQVDEVRRLVPEGTQKSPLEHLYVYVYNPGEVTRLTDRGYSHSCCSISQHCDFPISKY 572
DB 532 -----KGTQDTPLHHLIYVVSYEAAAGEIVRLTTPGFSHSCMSQNDFMEVSHY 579

RESULT 10
US-09-976-674-33
; Sequence 33, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-33

Query Match      60.0%; Score 2820.5; DB 2; Length 879;
Best Local Similarity 60.7%; Pred. No. 1.4e-278;
Matches 510; Conservative 132; Mismatches 183; Indels 15; Gaps 3;

QY 35 FVVERYSQKLLADTRKYHGYMMAKAPHDEMFKRNDPDGPHSDRIYYLAMSGENRE 94
DB 53 FQVKHSDGLRSIIHGSRYKSLGVNKAHPDFQVQKTDSESPHSHRLYYLGMYPGSR 112
QY 95 NTLFYSEIPKTNINRAAVALMSKPLDLFOATLDYMGYSREBELLRKRIGTVGIASYD 154
DB 113 NSLLYSEIPKPKYKREALLLSWKQMLDHFQATPHHGYISREBELLRKRIGTVGITSYD 172
QY 155 YHQSGLTFLQAGSGIVHVDGPGQGTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAF 214
DB 173 FHSSEGLFLFOANSLSLHCRDGGKNGFPMVSPMKPLRIKTCQSGPRMDPKICPADPAFFSF 232
QY 215 IHSNDIWSNTVREERLTYVHNELANMEEDARSAGVATFVLQEBFDYSGYWCCKAE 274
DB 233 INNSDLWANTETGEERLTFCHQGLSNVLDPPKSGAGVATFVIOEBFDRFTGYWCPTAS 292
QY 275 TTPSGG-KILRLIYEENDESEVEIIVHTVSPMLTTRADSPRYPKTGTANPKVTFKMSIM 333
DB 293 WEGSEGLKTLRLIYEENDESEVEIIVHVPSPALBERKTDVSYRPTGSKNPKIALKLBQ 352
QY 334 IDAEGRIIDVDKELIOPPEILFEGVEYIARAGWTPGKYAMSILLDRSOTRLQIVLISP 393
DB 353 TDSQKIVSTQEKLVQPFSSLPFKVEYIARAGWTRDKYAWAMFLDRPQWLQVLLPP 412
QY 394 ELFIPEVDVMERQRLIESVPDSTPLIYYEETDWINHIDIFHVFPQSH-EEETEFIF 452
DB 413 ALFIPSTENESORLASARAVPRNVQPVVYEEVTNWINVHIDIFYPFPQSEGEDELCLFLR 472
QY 453 ASECKTGRPHLYKITSLIKESKYRSSGGLPAPSDPKPIKEEIAITSGEWELGRHGSN 512
DB 473 ANECKTGCHLYKYTAVLKSQGYDWSPEFPFGDEDFKCPIKEEIAITSGEWELABHGS- 531
QY 513 IQVDEVRRLVPEGTQKSPLEHLYVYVYNPGEVTRLTDRGYSHSCCSISQHCDFPISKY 572
DB 532 -----KGTQDTPLHHLIYVVSYEAAAGEIVRLTTPGFSHSCMSQNDFMEVSHY 579
```



```
QY 498 ITSGEWELGRHSNIQVDEVRRLVYFEGTKDSPLFHLHYVSYVNPGEVRLTDRGYSH 557
Db 181 ITSGEWELGRHSNIQVDEVRRLVYFEGTKDSPLFHLHYVSYVNPGEVRLTDRGYSH 240
QY 558 SCCISQHCDFPISIKYNOXKPHCVSLYKLSPPDDPTCKTKFWMATILDSAGPLPDYTPP 617
Db 241 SCCISQHCDFPISIKYNOXKPHCVSLYKLSPPDDPTCKTKFWMATILDSAGPLPDYTPP 300
QY 618 EIFSPSTGTFTLGMLYKPHDLQPGKKYPTVLFYGGQVOLVNNRFGKVKYFRLNTLA 677
Db 301 EIFSPSTGTFTLGMLYKPHDLQPGKKYPTVLFYGGQ----- 340
QY 678 SLGYVVVVIDNRGSHRGLKPEGAFKYMGOIBDDQVEGLQYLASRYDFIDLDRVGIGH 737
Db 341 ----- 340
QY 738 WSYGGYLSLMMALMORSIDIFRAIAGAPVTLWIFDYTGTYERYMGHPDQNEQYILGVSAM 797
Db 341 -----VAIAGAPVTLWIFDYTGTYERYMGHPDQNEQYILGVSAM 380
QY 798 QAEKFPSPENRLLHGLFDENVHFHTSLLSFLVRAGKPYDLOIYPOERHSIRVPESG 857
Db 381 QAEKFPSPENRLLHGLFDENVHFHTSLLSFLVRAGKPYDLOIYPOERHSIRVPESG 440
QY 858 EHYELHLLHYLOENLGRSRIAAKVI 882
Db 441 EHYELHLLHYLOENLGRSRIAAKVI 465
```

```
RESULT 13
US-09-976-674-29
; Sequence 29, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-29
```

```
Query Match 51.2%; Score 2406; DB 2; Length 832;
Best Local Similarity 59.0%; Pred. No. 3.2e-236;
Matches 434; Conservative 122; Mismatches 177; Indels 2; Gaps 2;

QY 35 FVVERYSWQLKLLADTRKYHYMMAKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRE 94
Db 53 FQVQKHSWDGLRSIIHGSRKYSGLIVNKAHPDFQFVKTDGSPHSHRLYILGMPYGSRE 112
QY 95 NTLFYSEIPKTIINRAAIVMLSWKPLLDLFOATLDYGMYSREELLRERKRIGTVGIASYD 154
Db 113 NSLLYSEIPKPKVRKEALLLSWKQMLDHFQATPHHGVSREELLRERKRIGTVGITSYD 172
QY 155 YHOGSGTFLFOAGSGIYHVKGDPGQFTQQPLRPNLVETSCPNIRMDPKLCPADPDWIAF 214
Db 173 PHSESGLEFLFOASNSLFCRDKGNGKGMVSPKPLEIKTCQSPRMDPKICPADPAFTSF 232
QY 215 IHSNDIWSNTVTRERRLTVVHNELNMEEDARSAGVATFVLQEBFDRYSYWMCPKAE 274
Db 233 INNSDLWANIETGEERLLTFCHQGLSNVLDPKSAGVATFVQEBFDRFTGYMWCPTAS 292
```

```
QY 275 TTSPSG-KILRLTYEENDESEVEIHHVTSPMLTETRADSPRYPKTGTANPKVTFKMSIEM 333
Db 293 WEGSGELTKLRILEEVEDSEVEIHHVPSPALBERKTDTSYRPTGSKNPKIALKLAEPQ 352
QY 334 IDAERIIDVDIKELIQPFEILFEGVEYIARAGTWEGKYAMSIILLDRSQTRLOIVLISP 393
Db 353 TDSQKRIYSTQEKELVQPSLFFPKVYIARAGWRDGTAYAMWFLDRPOOMLQVLVLP 412
QY 394 ELFIPEVDDVMERQRLIESVPDSVTPLIYYEETDIIWINHDIHPVPPQSH-BEEIEFTF 452
Db 413 ALUFISTNEQRORLASARAVRNQPVYVVEEVNWINVHDIYFPFQSEGEDELCEFLR 472
QY 453 ASECTGRFRLHYKITSILSKYKRSGLPAPSPDFKCPIKEBIAITSGSEWELGRHGSN 512
Db 473 ANECTGCHLYKYTAVLKSQGYDMSBPPFGDEDFKCPIKEBIAITSGSEWELARHGSK 532
QY 513 IOVDEVRRLVYFEGTKDSPLFHLHYVSYVNPGEVRLTDRGYSHSCCSIQHCDFPISY 572
Db 533 IWNSETKLVTYFQGTCTPLEHLYVVSYEAAGEIVRLTTPGFSHSCSMNQNFDMFVSHY 592
QY 573 SNOKPHCVSLYKLSPPDDPTCKTKFWMATILDSAGPLPDYTPPEIFPSTGTFTLYG 632
Db 593 SSVSTPPCVHVYKLSGPDPLHKQPRFASMMEAASCPDPYVPPBIFPHFTRSDVRLY 652
QY 633 MLYKPHDLQPGKKYPTVLFYGGQVOLVNNRFGKVKYFRLNTLASLGYVVVVIDNRGSC 692
Db 653 MIYKPHALQPGKHTVLFYGGQVOLVNNRFGKIKYLRNLNTLASLGYAVVVIDGRSC 712
QY 693 HRGLPEGAFKYMGOIBDDQVEGLQYLASRYDFIDLDRVGIGHGYSGLYSLMALMQR 752
Db 713 QRGLRPEGALKNQMGQVEIQVEGLQFVAEKYFIDLRSVAIHGWSYGGFLSLMGLIHK 772
QY 753 SDIFRAIAGAPVTL 767
Db 773 PQVFRAQPLAYPPRL 787

RESULT 14
US-09-976-674-31
; Sequence 31, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-31

Query Match 51.2%; Score 2406; DB 2; Length 832;
Best Local Similarity 59.0%; Pred. No. 3.2e-236;
Matches 434; Conservative 122; Mismatches 177; Indels 2; Gaps 2;

QY 35 FVVERYSWQLKLLADTRKYHYMMAKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRE 94
Db 53 FQVQKHSWDGLRSIIHGSRKYSGLIVNKAHPDFQFVKTDGSPHSHRLYILGMPYGSRE 112
QY 95 NTLFYSEIPKTIINRAAIVMLSWKPLLDLFOATLDYGMYSREELLRERKRIGTVGIASYD 154
Db 113 NSLLYSEIPKPKVRKEALLLSWKQMLDHFQATPHHGVSREELLRERKRIGTVGITSYD 172
QY 155 YHOGSGTFLFOAGSGIYHVKGDPGQFTQQPLRPNLVETSCPNIRMDPKLCPADPDWIAF 214
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 13:01:45 ; Search time 172 Seconds
(without alignments)
2142.590 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAAMETEQLGVETFTADC.....HLHLVQLQENLGSRIAALKVI 882

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main.*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pcp.*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pcp.*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pcp.*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pcp.*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4700	100.0	882	3	US-09-976-674-1
2	4700	100.0	882	4	US-10-054-776-2
3	4700	100.0	882	4	US-10-170-789-38
4	4700	100.0	882	4	US-10-311-035-9
5	4700	100.0	882	4	US-10-072-012-622
6	4700	100.0	882	4	US-10-415-122-6
7	4700	100.0	882	4	US-10-825-632-1
8	4700	100.0	882	5	US-10-982-512-1
9	4528.5	96.4	883	4	US-10-072-012-621
10	3504	74.6	658	3	US-09-976-674-19
11	3504	74.6	658	5	US-10-982-512-19
12	3504	74.6	661	3	US-09-976-674-11
13	3504	74.6	661	5	US-10-982-512-11
14	3504	74.6	690	3	US-09-976-674-7
15	3504	74.6	690	5	US-10-982-512-7
16	3236	68.9	613	3	US-09-976-674-21
17	3236	68.9	613	5	US-10-982-512-21
18	2870	61.1	863	3	US-09-976-674-3
19	2870	61.1	863	4	US-10-072-012-619
20	2870	61.1	863	5	US-10-982-512-3
21	2870	61.1	892	3	US-09-976-674-23
22	2870	61.1	892	3	US-09-976-674-27
23	2870	61.1	892	5	US-10-982-512-23
24	2870	61.1	892	5	US-10-982-512-27
25	2870	61.1	892	5	US-10-433-757-12
26	2863	60.9	969	4	US-10-415-122-2
27	2862	60.9	863	4	US-10-072-012-224

28	2862	60.9	863	4	US-10-072-012-226	Sequence 226, Appl
29	2835	60.3	830	4	US-10-415-122-7	Sequence 7, Appl
30	2833	60.3	869	4	US-10-415-122-4	Sequence 4, Appl
31	2820.5	60.0	879	3	US-09-976-674-33	Sequence 33, Appl
32	2820.5	60.0	879	3	US-09-976-674-35	Sequence 35, Appl
33	2820.5	60.0	879	5	US-10-982-512-33	Sequence 33, Appl
34	2820.5	60.0	879	5	US-10-982-512-35	Sequence 35, Appl
35	2547.5	54.2	580	6	US-10-275-505-2	Sequence 2, Appl
36	2547.5	54.2	580	6	US-11-140-224-2	Sequence 2, Appl
37	2422	51.5	465	4	US-10-825-632-5	Sequence 5, Appl
38	2406	51.2	832	3	US-09-976-674-29	Sequence 29, Appl
39	2406	51.2	832	3	US-09-976-674-31	Sequence 31, Appl
40	2406	51.2	832	5	US-10-982-512-29	Sequence 29, Appl
41	2406	51.2	832	5	US-10-982-512-31	Sequence 31, Appl
42	2397	51.0	689	4	US-10-072-012-620	Sequence 620, App
43	2356.5	50.1	819	3	US-09-976-674-37	Sequence 37, Appl
44	2356.5	50.1	819	3	US-09-976-674-39	Sequence 39, Appl
45	2356.5	50.1	819	5	US-10-982-512-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Query Match	100.0%	Score 4700;	DB 3;	Length 882;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches	882;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MAAAMETEQLGVETFTADCENIESQDRPKLEPPFVVERYSQKLLADTRKYHYGM 60		
DB	1	MAAAMETEQLGVETFTADCENIESQDRPKLEPPFVVERYSQKLLADTRKYHYGM 60		
QY	61	AKAHPDMFKVRNDPDPGPHSDRIYYLAMSGENRENTLFYSEIPKTNRAAVALMSWKPLL 120		
DB	61	AKAHPDMFKVRNDPDPGPHSDRIYYLAMSGENRENTLFYSEIPKTNRAAVALMSWKPLL 120		
QY	121	DLFOATLDYGMYSREELLERKRI GTVGIASYDVHQSGTFLFOAGSGIYHVKGQGG 180		
DB	121	DLFOATLDYGMYSREELLERKRI GTVGIASYDVHQSGTFLFOAGSGIYHVKGQGG 180		
QY	181	FTQQLPRLNVLVETSCPNIRMDPKLPADPDWIAFIHSNDIWSNIVTREERLTYYVHNL 240		
DB	181	FTQQLPRLNVLVETSCPNIRMDPKLPADPDWIAFIHSNDIWSNIVTREERLTYYVHNL 240		
QY	241	ANMBEDARSAGATVFLQEEFDYSGVWCPKATTPSGGKILRIIYENDESEVEI IHV 300		
DB	241	ANMBEDARSAGATVFLQEEFDYSGVWCPKATTPSGGKILRIIYENDESEVEI IHV 300		
QY	301	TSPMLTERRADSFYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIOPFEILFEGVE 360		
DB	301	TSPMLTERRADSFYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIOPFEILFEGVE 360		

```
QY 361 YIARAGWTEGKYAWSILLDRSQTQLQIVLISPELFTPVEDDVMERQRLIESVPSVTPL 420
Db 361 YIARAGWTEGKYAWSILLDRSQTQLQIVLISPELFTPVEDDVMERQRLIESVPSVTPL 420
QY 421 IYEEETDIIWINIHDIHFVFPQSHEEBIEIFASECTGPRHLYKITSLKESKYRSG 480
Db 421 IYEEETDIIWINIHDIHFVFPQSHEEBIEIFASECTGPRHLYKITSLKESKYRSG 480
QY 481 GLPAPSDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLYVVS 540
QY 541 YNPGEVTRLDRGYSHSCCISQHCDFFIISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db 541 YNPGEVTRLDRGYSHSCCISQHCDFFIISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFTYGGPQVQL 660
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFTYGGPQVQL 660
QY 661 VNNRFGVKYFRNLNTLASLGYVVVVVINDRGSCHRGKLFEGAFYKMGQIEIDDOVEGLQY 720
Db 661 VNNRFGVKYFRNLNTLASLGYVVVVVINDRGSCHRGKLFEGAFYKMGQIEIDDOVEGLQY 720
QY 721 LASRYDFDLDVRGIGHWSYGGYLSLALMQRSDIFRVAITAGAPVTLWIFYDTGYTERYM 780
Db 721 LASRYDFDLDVRGIGHWSYGGYLSLALMQRSDIFRVAITAGAPVTLWIFYDTGYTERYM 780
QY 781 GHPDQNEQGYLGVSVMQAEKFPSEPNRLLHLLHGFLENVHFAHTSILLSPFLVRAGKPYD 840
Db 781 GHPDQNEQGYLGVSVMQAEKFPSEPNRLLHLLHGFLENVHFAHTSILLSPFLVRAGKPYD 840
QY 841 LQIYQERHSIRVPESGHEHLLHYLQENLGSRTAALKVI 882
Db 841 LQIYQERHSIRVPESGHEHLLHYLQENLGSRTAALKVI 882
```

RESULT 2

```
US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US20030165618A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2
```

```
Query Match
Best Local Similarity 100.0%; Score 4700; DB 4; Length 882;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAMETEQLGVEIFETADCEENTIESQDRPKLPFFVVERYSQOLKLLADTRKHGYMM 60
Db 1 MAAMETEQLGVEIFETADCEENTIESQDRPKLPFFVVERYSQOLKLLADTRKHGYMM 60
QY 61 AKAPHDFVVRNDPDGPHSDRIYLLAMSGENRENTLYSEIPIKTNINRAVLMLSWKPLL 120
Db 61 AKAPHDFVVRNDPDGPHSDRIYLLAMSGENRENTLYSEIPIKTNINRAVLMLSWKPLL 120
QY 121 DLFOATLDYGMYSREELLRRKRIGTVGTIGASVDYHGGSGTFLFQAGSGIYHVKGQGG 180
Db 121 DLFOATLDYGMYSREELLRRKRIGTVGTIGASVDYHGGSGTFLFQAGSGIYHVKGQGG 180
```

```
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWI SNIVTREERRLTYYVHNL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWI SNIVTREERRLTYYVHNL 240
QY 241 ANNEEDASAGVATVFLQEBEDRYSGYWWCPKAETTPSGGKILRILYEENDESEVEI IHV 300
Db 241 ANNEEDASAGVATVFLQEBEDRYSGYWWCPKAETTPSGGKILRILYEENDESEVEI IHV 300
QY 301 TSPMLETRADSFRIYPTKTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIQPFILPEGVE 360
Db 301 TSPMLETRADSFRIYPTKTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIQPFILPEGVE 360
QY 361 YIARAGWTEGKYAWSILLDRSQTQLQIVLISPELFTPVEDDVMERQRLIESVPSVTPL 420
Db 361 YIARAGWTEGKYAWSILLDRSQTQLQIVLISPELFTPVEDDVMERQRLIESVPSVTPL 420
QY 421 IYEEETDIIWINIHDIHFVFPQSHEEBIEIFASECTGPRHLYKITSLKESKYRSG 480
Db 421 IYEEETDIIWINIHDIHFVFPQSHEEBIEIFASECTGPRHLYKITSLKESKYRSG 480
QY 481 GLPAPSDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLYVVS 540
QY 541 YNPGEVTRLDRGYSHSCCISQHCDFFIISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db 541 YNPGEVTRLDRGYSHSCCISQHCDFFIISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVQL 660
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVQL 660
QY 661 VNNRFGVKYFRNLNTLASLGYVVVVVINDRGSCHRGKLFEGAFYKMGQIEIDDOVEGLQY 720
Db 661 VNNRFGVKYFRNLNTLASLGYVVVVVINDRGSCHRGKLFEGAFYKMGQIEIDDOVEGLQY 720
QY 721 LASRYDFDLDVRGIGHWSYGGYLSLALMQRSDIFRVAITAGAPVTLWIFYDTGYTERYM 780
Db 721 LASRYDFDLDVRGIGHWSYGGYLSLALMQRSDIFRVAITAGAPVTLWIFYDTGYTERYM 780
QY 781 GHPDQNEQGYLGVSVMQAEKFPSEPNRLLHLLHGFLENVHFAHTSILLSPFLVRAGKPYD 840
Db 781 GHPDQNEQGYLGVSVMQAEKFPSEPNRLLHLLHGFLENVHFAHTSILLSPFLVRAGKPYD 840
QY 841 LQIYQERHSIRVPESGHEHLLHYLQENLGSRTAALKVI 882
Db 841 LQIYQERHSIRVPESGHEHLLHYLQENLGSRTAALKVI 882
```

RESULT 3

```
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCI/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-170-789-38

Query Match      100.0%; Score 4700; DB 4; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAAAMETEOLGVEIFETADCEENIESQDRPKLEPPYVERYSWSOLKLLADTKYHYNM 60
Db      1  MAAAMETEOLGVEIFETADCEENIESQDRPKLEPPYVERYSWSOLKLLADTKYHYNM 60

Qy      61  AKAPHDFMFVKXNDPDGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMSWKPLL 120
Db      61  AKAPHDFMFVKXNDPDGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMSWKPLL 120

Qy      121  DLFOATLDYGMYSREELLREKRIGTVGSIADYHOGSGTFLPQAGSGIYHVKGDPQG 180
Db      121  DLFOATLDYGMYSREELLREKRIGTVGSIADYHOGSGTFLPQAGSGIYHVKGDPQG 180

Qy      181  FTQOPLRPNLVETSCFNIRMDPKLPADPDWIAFTHSNDIWSINIVTREERLTYVHNL 240
Db      181  FTQOPLRPNLVETSCFNIRMDPKLPADPDWIAFTHSNDIWSINIVTREERLTYVHNL 240

Qy      241  ANNEEDARSAGVATFVLQEEFDRYSGYWCPCAETTTSGGKILRLIYEENDESEVEIHV 300
```

```

Db      241  ANNEEDARSAGVATFVLQEEFDRYSGYWCPCAETTTSGGKILRLIYEENDESEVEIHV 300
Qy      301  TSPMLETRRADSFYRYPKGTGTANPKVTFKQSEIMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db      301  TSPMLETRRADSFYRYPKGTGTANPKVTFKQSEIMIDAEGRIIDVIDKELIQPFELFEGVE 360
Qy      361  YIARAGWTPEGKYAWSILLDRSQRLQIVLSPELFPVEDDVNERQLLESVPDSVTPL 420
Db      361  YIARAGWTPEGKYAWSILLDRSQRLQIVLSPELFPVEDDVNERQLLESVPDSVTPL 420
Qy      421  IYEEETDIIWINHDIHFVPOSHEEIEFIPASECKTGRRLHYKITSILKESKYKSSG 480
Db      421  IYEEETDIIWINHDIHFVPOSHEEIEFIPASECKTGRRLHYKITSILKESKYKSSG 480
Qy      481  GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLLYVS 540
Db      481  GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLLYVS 540
Qy      541  YNPGVETRLTDRGYSHSCCISQHCDFPISKYSQKQNPCHCVSLYKLSPEDDPTCKTKEF 600
Db      541  YNPGVETRLTDRGYSHSCCISQHCDFPISKYSQKQNPCHCVSLYKLSPEDDPTCKTKEF 600
Qy      601  WATILDSAGPLPDYTPPEIFSPESSTGFTLYGMLYKPHDLQPGKKYPTVLFYGGPOVL 660
Db      601  WATILDSAGPLPDYTPPEIFSPESSTGFTLYGMLYKPHDLQPGKKYPTVLFYGGPOVL 660
Qy      661  VNNRFGVKYFRNLTLASLGYYVVVIDNRRGSHRGKPEGAFKYMKGQIEIDDOVEGLQY 720
Db      661  VNNRFGVKYFRNLTLASLGYYVVVIDNRRGSHRGKPEGAFKYMKGQIEIDDOVEGLQY 720
Qy      721  LASRYDFIDLDRVGIHGSYGYLSLMMALMORSIDIFRVAIAGAPVTLMIFDTGYTRYM 780
Db      721  LASRYDFIDLDRVGIHGSYGYLSLMMALMORSIDIFRVAIAGAPVTLMIFDTGYTRYM 780
Qy      781  GHPDQNEQGYLGSVAMQAEKPPSEPNRLILHGLDENVHFAHTSILLSPVLRAGPYD 840
Db      781  GHPDQNEQGYLGSVAMQAEKPPSEPNRLILHGLDENVHFAHTSILLSPVLRAGPYD 840
Qy      841  LQIYPQERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882
Db      841  LQIYPQERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882

RESULT 4
US-10-311-035-9
; Sequence 9, Application US/10311035
; Publication No. US2004002343A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAU, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOLEY, Catherine M.
; APPLICANT: DEBEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Darniel B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAPALIA, April
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyrung Aina M.
; APPLICANT: ARVIZO, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUNAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
```

APPLICANT: DAS, Depopriya
APPLICANT: KEARNEY, Liam
APPLICANT: KALLICK, Deborah A.
TITLE OF INVENTION: Proteases
FILE REFERENCE: PI-0123 PCT
CURRENT APPLICATION NUMBER: US/10/311.035
PRIOR FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 882
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
US-10-311-035-9

Query Match 100.0%; Score 4700; DB 4; Length 882;
Best Local Similarity 100.0%; Pred No 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAMETEOLGVIEFETADCFENIESQDRPKLEPPYVERYSWSQLKLLADTRKHGYMM 60
Db 1 MAAMETEOLGVIEFETADCFENIESQDRPKLEPPYVERYSWSQLKLLADTRKHGYMM 60
Qy 61 AKAPHDFMFKVNDPDGPHSDRIYYILAMSGENRENTLFYSEIPIKTNRAAVLMSWKPLL 120
Db 61 AKAPHDFMFKVNDPDGPHSDRIYYILAMSGENRENTLFYSEIPIKTNRAAVLMSWKPLL 120
Qy 121 DLFOATFLDYGMYSREELLLRRKRGITGVGIASYDYGSGTFLFQAGSGIYHVKGPGOG 180
Db 121 DLFOATFLDYGMYSREELLLRRKRGITGVGIASYDYGSGTFLFQAGSGIYHVKGPGOG 180
Qy 181 FTQOPLRPNLVETSCFNIRMDPKCPADPDWIAPIHNSDIWISNIVTREERLTYVHNEL 240
Db 181 FTQOPLRPNLVETSCFNIRMDPKCPADPDWIAPIHNSDIWISNIVTREERLTYVHNEL 240
Qy 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHV 300
Db 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHV 300
Qy 301 TSPMLETRRADSFYRYPKGTANPKVTFKMSIIMIDAGRIIDVIDKELIQPPEILPEGYE 360
Db 301 TSPMLETRRADSFYRYPKGTANPKVTFKMSIIMIDAGRIIDVIDKELIQPPEILPEGYE 360
Qy 361 YIARAGWTPEGKYANSILLDRSOTLQIVLISPELFIPEVDDVNERQLIESVPDSVTPL 420
Db 361 YIARAGWTPEGKYANSILLDRSOTLQIVLISPELFIPEVDDVNERQLIESVPDSVTPL 420
Qy 421 IYEEETDIIWNIHDIHFVFPQSHEEIEFTIPASECKTGFRHLYKITSILKSKYKRSSG 480
Db 421 IYEEETDIIWNIHDIHFVFPQSHEEIEFTIPASECKTGFRHLYKITSILKSKYKRSSG 480
Qy 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVS 540
Db 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVS 540
Qy 541 YVNPGEVTRTLDRGSHSHSCCISQHCDFPISKYSNQKNPCHCVSLYKLSPPDDDTCKTKEF 600
Db 541 YVNPGEVTRTLDRGSHSHSCCISQHCDFPISKYSNQKNPCHCVSLYKLSPPDDDTCKTKEF 600
Qy 601 WATILDSAGLPDYTPPEIFSPFESTTGTFLYGLMKYKPHDLQPGKKYPTVLIYGGPQVQL 660
Db 601 WATILDSAGLPDYTPPEIFSPFESTTGTFLYGLMKYKPHDLQPGKKYPTVLIYGGPQVQL 660
Qy 661 VNNRFKGVKYFRNLTLASLGYYVVVVIDNRGSHRGLKPEGAFYKMGQIIDDQVEGLQY 720
Db 661 VNNRFKGVKYFRNLTLASLGYYVVVVIDNRGSHRGLKPEGAFYKMGQIIDDQVEGLQY 720
Qy 721 LASRYDFIDLDVRVGIHGSYGGYLSLMALMORSIDIFRVAIAGAPVTLWIFDYDGYTERYM 780

Db 721 LASRYDFIDLDVRVGIHGSYGGYLSLMALMORSIDIFRVAIAGAPVTLWIFDYDGYTERYM 780
Qy 781 GHPDNEQGYVLGSVMAQAEKFPSEPNRLLLHGHCFLDENVHFAHTSILLSFVLRACKPYD 840
Db 781 GHPDNEQGYVLGSVMAQAEKFPSEPNRLLLHGHCFLDENVHFAHTSILLSFVLRACKPYD 840
Qy 841 LQIYQERHSIRVPESGHEHYELHLLHYLOENLGSRIAALKVI 882
Db 841 LQIYQERHSIRVPESGHEHYELHLLHYLOENLGSRIAALKVI 882

RESULT 5

US-10-072-012-622
Sequence 622, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shinkete, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigar, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier, Jr. Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosbe, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072.012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 622
LENGTH: 882
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-622

; CURRENT APPLICATION NUMBER: US/10/415,122
 CURRENT FILING DATE: 2003-08-07

CURRENT FILING DATE: 2003-03-01
NUMBER OF SEC ID NOS: 8

; NUMBER OF SEQ ID NOS: 6
CONTENTS: Defect In Warg

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 6

; LENGTH: 882

TYPE: PRT

ORGANISM: *Homo sapiens*

US-10-415-122-6

Query Match 100.0%; Score 4700; DB 4; Length 882;

Query Match
best recall similarity 100.0% Pred. No: 0;
best recall similarity 100.0% Pred. No: 0;

Best Local Similarity	100.0%, Prec. no. 87
Mismatches	0; Indels 0

QY	1	MAAA	ME	TE	Q	L	G	V	E	I	F	T	A	C	E	E	N	I	S	O	D	R	P	K	L	E	P	F	V	E	R	Y	S	W	S	Q	L	K	L	L	A	D	T	R	K	Y	H	G	M	60												
Db	1	MAAA	ME	TE	Q	L	G	V	E	I	F	T	A	C	E	E	N	I	S	O	D	R	P	K	L	E	P	F	V	E	R	Y	S	W	S	Q	L	K	L	L	A	D	T	R	K	Y	H	G	M	60												
QY	61	AKA	PH	D	F	M	F	V	K	R	N	D	P	G	H	S	D	R	I	Y	I	L	A	M	S	G	E	N	T	L	F	S	E	I	P	K	T	I	N	R	A	A	V	L	M	L	S	W	K	P	L	120										
Db	61	AKA	PH	D	F	M	F	V	K	R	N	D	P	G	H	S	D	R	I	Y	I	L	A	M	S	G	E	N	T	L	F	S	E	I	P	K	T	I	N	R	A	A	V	L	M	L	S	W	K	P	L	120										
QY	121	DL	F	O	A	T	L	D	Y	G	M	Y	S	R	E	E	L	L	R	K	R	I	G	T	V	G	I	A	S	D	Y	H	Q	S	G	T	F	L	F	O	A	G	S	G	I	H	V	K	D	G	P	Q	180									
Db	121	DL	F	O	A	T	L	D	Y	G	M	Y	S	R	E	E	L	L	R	K	R	I	G	T	V	G	I	A	S	D	Y	H	Q	S	G	T	F	L	F	O	A	G	S	G	I	H	V	K	D	G	P	Q	180									
QY	181	F	T	Q	O	P	L	R	N	L	V	E	T	S	C	P	N	I	R	M	D	P	K	L	C	A	D	P	O	W	I	A	F	I	H	S	N	D	I	W	I	S	N	I	V	T	R	E	R	L	T	V	N	H	E	L	240					
Db	181	F	T	Q	O	P	L	R	N	L	V	E	T	S	C	P	N	I	R	M	D	P	K	L	C	A	D	P	O	W	I	A	F	I	H	S	N	D	I	W	I	S	N	I	V	T	R	E	R	L	T	V	N	H	E	L	240					
QY	241	AN	ME	E	D	A	R	S	A	G	V	A	T	F	L	O	E	F	D	R	Y	S	G	Y	W	M	C	P	K	A	E	T	T	P	S	G	K	L	R	I	L	I	Y	E	N	D	E	S	E	V	E	I	I	H	V	300						
Db	241	AN	ME	E	D	A	R	S	A	G	V	A	T	F	L	O	E	F	D	R	Y	S	G	Y	W	M	C	P	K	A	E	T	T	P	S	G	K	L	R	I	L	I	Y	E	N	D	E	S	E	V	E	I	I	H	V	300						
QY	301	T	S	P	M	L	E	T	R	A	D	S	F	R	Y	P	K	T	G	T	A	N	P	K	V	T	F	K	M	S	E	I	M	I	D	A	E	G	R	I	I	D	V	I	D	K	E	L	L	O	P	E	I	L	F	E	G	V	E	360		
Db	301	T	S	P	M	L	E	T	R	A	D	S	F	R	Y	P	K	T	G	T	A	N	P	K	V	T	F	K	M	S	E	I	M	I	D	A	E	G	R	I	I	D	V	I	D	K	E	L	L	O	P	E	I	L	F	E	G	V	E	360		
QY	361	Y	I	A	R	A	G	W	T	P	E	G	K	Y	A	W	S	I	L	L	D	R	S	O	T	R	L	O	I	V	L	I	S	P	E	L	F	I	P	V	E	D	D	Y	M	E	R	Q	R	L	I	E	S	V	P	D	S	V	T	P	L	420
Db	361	Y	I	A	R	A	G	W	T	P	E	G	K	Y	A	W	S	I	L	L	D	R	S	O	T	R	L	O	I	V	L	I	S	P	E	L	F	I	P	V	E	D	D																			

RESIT.T 6

RESULTS
IIS-10-415-122-6

US-10-413-122-8 Application US/10415122

; sequence 6, Application US/1971-00053369A1

; Publication No. US200

; GENERAL INFORMATION:

APPLICANT: THE UNIVERSITY OF SYDNEY

TITLE OF INVENTION: DIPH

```
RESULT 7
US-10-825-632-1
; Sequence 1, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-1

Query Match      100.0%; Score 4700; DB 4; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVIFETADCEENIESQDRPKLEPPFVVERYSWSQKLLADTRKXHYGMM 60
DB 1 MAAAMETEQLGVIFETADCEENIESQDRPKLEPPFVVERYSWSQKLLADTRKXHYGMM 60
QY 61 AKAPHPDMFVKRNDPDGPHSDRIYGLAMSGENRENTLFYSEIPKTIINRAAVLMSWPKLL 120
DB 61 AKAPHPDMFVKRNDPDGPHSDRIYGLAMSGENRENTLFYSEIPKTIINRAAVLMSWPKLL 120
QY 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLYVHNEL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLYVHNEL 240
QY 241 ANNEEDARSAGVATFVLOEEFDYSGYVWCPKATTPSGGKILRIIYEENDESEVEIHW 300
DB 241 ANNEEDARSAGVATFVLOEEFDYSGYVWCPKATTPSGGKILRIIYEENDESEVEIHW 300
QY 301 TSPMLETRADSPRYPKTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIQPFILFEGVE 360
DB 301 TSPMLETRADSPRYPKTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIQPFILFEGVE 360

RESULT 8
US-10-982-512-1
; Sequence 1, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Oi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Juichen, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-1

Query Match      100.0%; Score 4700; DB 5; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVIFETADCEENIESQDRPKLEPPFVVERYSWSQKLLADTRKXHYGMM 60
DB 1 MAAAMETEQLGVIFETADCEENIESQDRPKLEPPFVVERYSWSQKLLADTRKXHYGMM 60
QY 61 AKAPHPDMFVKRNDPDGPHSDRIYGLAMSGENRENTLFYSEIPKTIINRAAVLMSWPKLL 120
DB 61 AKAPHPDMFVKRNDPDGPHSDRIYGLAMSGENRENTLFYSEIPKTIINRAAVLMSWPKLL 120
QY 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLYVHNEL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLYVHNEL 240
QY 241 ANNEEDARSAGVATFVLOEEFDYSGYVWCPKATTPSGGKILRIIYEENDESEVEIHW 300
DB 241 ANNEEDARSAGVATFVLOEEFDYSGYVWCPKATTPSGGKILRIIYEENDESEVEIHW 300
QY 301 TSPMLETRADSPRYPKTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIQPFILFEGVE 360
DB 301 TSPMLETRADSPRYPKTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIQPFILFEGVE 360
```

QY	361	YIARAGWTPEGKYAWSIILLDRSOTRLQVLVISPFLIPVEDDVMQRQLIESVPDSVTP	420
Db	361	YIARAGWTPEGKYAWSIILLDRSOTRLQVLVISPFLIPVEDDVMQRQLIESVPDSVTP	420
QY	421	LIYEETDIWINIHDI FHVFPQSHEEIEFIPASECTGPRHLKYKITSILKESKYKSSG	480
Db	421	LIYEETDIWINIHDI FHVFPQSHEEIEFIPASECTGPRHLKYKITSILKESKYKSSG	480
QY	481	GLPAPSDPKPIKEBIATTSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS	540
Db	481	GLPAPSDPKPIKEBIATTSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS	540
QY	541	YVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF	600
Db	541	YVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF	600
QY	601	FWATILDSAGPLPDYTPPIFSPFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVQ	660
Db	601	FWATILDSAGPLPDYTPPIFSPFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVQ	660
QY	661	VNNRFGVKYFRLNTLASLGYVVVVIDNRGSGCHRLKFEPAFKYKMGQIEIDQVEGLQY	720
Db	661	VNNRFGVKYFRLNTLASLGYVVVVIDNRGSGCHRLKFEPAFKYKMGQIEIDQVEGLQY	720
QY	721	LASRYDFIDLRVGIHGSYGYLSLMALMORSIDIFRVAIAGAPVTILWIFDYTGTYTERYM	780
Db	721	LASRYDFIDLRVGIHGSYGYLSLMALMORSIDIFRVAIAGAPVTILWIFDYTGTYTERYM	780
QY	781	GHPDQNEQGYLYGSVAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSLFLVRACKPYD	840
Db	781	GHPDQNEQGYLYGSVAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSLFLVRACKPYD	840
QY	841	LQIYPOERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI	882
Db	841	LQIYPOERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI	882
RESULT 9			
US-10-072-012-621			
; Sequence 621, Application US/10072012			
; Publication No. US20040033493A1			
; GENERAL INFORMATION:			
; APPLICANT: Tchernev, Velizar			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Zernhusen, Bryan			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Shimkets, Richard			
; APPLICANT: Li, Li			
; APPLICANT: Gangolli, Esha			
; APPLICANT: Padigar, Muralidhara			
; APPLICANT: Anderson, David W.			
; APPLICANT: Rastelli, Luca			
; APPLICANT: Miller, Charles E.			
; APPLICANT: Gerlach, Valerie			
; APPLICANT: Taupier Jr, Raymond J.			
; APPLICANT: Gusev, Vladimiri Y.			
; APPLICANT: Colman, Steven D.			
; APPLICANT: Wolenc, Adam R.			
; APPLICANT: Pena, Carol E. A			
; APPLICANT: Furtak, Katarzyna			
; APPLICANT: Grosse, William M.			
; APPLICANT: Alsobrook II, John P.			
; APPLICANT: Lepley, Denise M.			
; APPLICANT: Rieger, Daniel K.			
; APPLICANT: Burgess, Catherine E.			
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
; FILE REFERENCE: 21402-258			
; CURRENT APPLICATION NUMBER: US/10/072,012			
; PRIOR FILING DATE: 2002-01-31			
; CURRENT APPLICATION NUMBER: 60/265,102			
; PRIOR FILING DATE: 2001-01-30			
; PRIOR APPLICATION NUMBER: 60/265,514			

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,517

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,412

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,395

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/266,406

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 60/266,767

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/267,057

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/266,975

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/267,459

PRIOR FILING DATE: 2001-02-08

PRIOR APPLICATION NUMBER: 60/267,459

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1391

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 621

LENGTH: 883

TYPE: PRT

ORGANISM: Mus musculus

US-10-072-012-621

Query Match 96.4%; Score 4528.5; DB 4; Length 883;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 845; Conservative 18; Mismatches 19; Indels 1; Gaps 1;

QY	1	MAAAMETEOLGVEIPEITACBEN-NIESODRPKLEPFYVERYSWOLKLLADTRKYHYM	59
Db	1	MAAAMETEOLGVEIPEITACBEN-NIESODRPKLEPFYVERYSWOLKLLADTRKYHYM	60
QY	60	MAKAPHDFMFKVENDPDGPHSDRIYVYLAWSNGENRENTLIFYSEIPIKTINRAAVLMSWKPL	119
Db	61	MAKAPHDFMFKVENDPDGPHSDRIYVYLAWSNGENRENTLIFYSEIPIKTINRAAVLMSWKPL	120
QY	120	LDLFQATLDYGMYSREELRLRKRITGVGTASDYHOGSGTFLFQAGSGIYHVKGSGPQ	179
Db	121	LDLFQATLDYGMYSREELRLRKRITGVGTAAADYHOGSGTFLFQAGSGIYHVKGSGPH	180
QY	180	GFTQQPLRPNLVETSCPNIRMDPKCPADPDWIAFHNSNDIWNISNIVTRERRRLTYVHNE	239
Db	181	GFTQQPLRPNLVETSCPNIRMDPKCPADPDWIAFHNSNDIWNISNIVTRERRRLTYVHNE	240
QY	240	LANNEEDARSAGVATVFLQEEEDRYSGYVWMCPCAKETTPSGGKILRLIYEENDESEVELIH	299
Db	241	LANNEEDARSAGVATVFLQEEEDRYSGYVWMCPCAKETTPSGGKILRLIYEENDESEVELIH	300
QY	300	VTSPLMLETTRADSFYRYPKTGTANPKVTPKMSIEMIDAEGRIIDVIDKELIQPFELFEGV	359
Db	301	VTSPLMLETTRADSFYRYPKTGTANPKVTPKMSIEMIDAEGRIIDVIDKELIQPFELFEGV	360
QY	360	EYIARAGWTPEGKYAWSIILLDRSOTRLQVLVISPFLIPVEDDVMQRQLIESVPDSVTP	419
Db	361	EYIARAGWTPEGKYAWSIILLDRSOTRLQVLVISPFLIPVEDDVMQRQLIESVPDSVTP	420
QY	420	LIYEETDIWINIHDI FHVFPQSHEEIEFIPASECTGPRHLKYKITSILKESKYKSS	479
Db	421	LIYEETDIWINIHDI FHVFPQSHEEIEFIPASECTGPRHLKYKITSILKESKYKSS	480
QY	480	GLPAPSDPKPIKEBIATTSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS	539
Db	481	GLPAPSDPKPIKEBIATTSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS	540
QY	540	SYVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKE	599
Db	541	SYVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKE	600
QY	600	FWATILDSAGPLPDYTPPIFSPFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVQ	659
Db	601	FWATILDSAGPLPDYTPPIFSPFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVQ	660

QY 660 LVNNRFGKGVYFRNLNTLASGYVVVVVDNRGSHRGGLKPEGAPKYKMGQIEIDDOVEGLQ 719
DB 661 LVNNRFGKGVYFRNLNTLASGYVVVVVDNRGSHRGGLKPEGAPKYKMGQIEIDDOVEGLQ 720
QY 720 YLASRYDFIDLDVRGIHGSYGGYLSLMLMQRSDIPRVAIAGAPVTLWIFDYDTGYTERY 779
DB 721 YLASQYDFIDLDVRGIHGSYGGYLSLMLMQRSDIPRVAIAGAPVTLWIFDYDTGYTERY 780
QY 780 MGHPPQNEGGYVLSGVAQAEKFPSEPNRLILLHGFLDENVHFAHTSILLSLFLVRAGKPY 839
DB 781 MGHPPQNEGGYVLSGVAQAEKFPSEPNRLILLHGFLDENVHFAHTSILLSLFLVRAGKPY 840
QY 840 DLQIYPOERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882
DB 841 DLQIYPOERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 883

RESULT 10

US-09-976-674-19
; Sequence 19, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Query Match 74.6%; Score 3504; DB 3; Length 658;
Best Local Similarity 100.0%; Pred. No. 7.3e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMETEOLGVEIFETADCEENTIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
DB 1 MAAAMETEOLGVEIFETADCEENTIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
QY 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSNGENRNTLFYSEI PKTINRAAVLMLSWKPLL 120
DB 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSNGENRNTLFYSEI PKTINRAAVLMLSWKPLL 120
QY 121 DLFOATLDYGMYSRREELLRRKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSRREELLRRKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVTSCPNIRMDPKLCPADPDWIAFIHSNDIWSNIVTREERLTYVHNEL 240
DB 181 FTQOPLRPNLVTSCPNIRMDPKLCPADPDWIAFIHSNDIWSNIVTREERLTYVHNEL 240
QY 241 ANNEEDARSAGVATFVLOEEDFDRYSGYWCPCAEKTPSPGGKILRIIYEENDESEVEI IHV 300
DB 241 ANNEEDARSAGVATFVLOEEDFDRYSGYWCPCAEKTPSPGGKILRIIYEENDESEVEI IHV 300
QY 301 TSPMLTRADSFYPKTGTPANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
DB 301 TSPMLTRADSFYPKTGTPANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
QY 361 YIARAGWTPEGKYANSLILDRSOTRLQIVLISPELFIPEVDVNMERORLIESVPDSVTPL 420
DB 361 YIARAGWTPEGKYANSLILDRSOTRLQIVLISPELFIPEVDVNMERORLIESVPDSVTPL 420

QY 421 LIYBETTDIWINIHDIHFVPOSHERRIERIPASECKTGRPHLYKITSLIKESKYKRSSG 480
DB 421 LIYBETTDIWINIHDIHFVPOSHERRIERIPASECKTGRPHLYKITSLIKESKYKRSSG 480
QY 481 GLPAPSDPKPIKEBIATISGEWEVLGRHGSNTQVDEVRLVYFEGTKDSPLEHHLYVVS 540
DB 481 GLPAPSDPKPIKEBIATISGEWEVLGRHGSNTQVDEVRLVYFEGTKDSPLEHHLYVVS 540
QY 541 YNPGGEVTRLDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKGF 600
DB 541 YNPGGEVTRLDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKGF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKPYTVLFIYGG 655
DB 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKPYTVLFIYGG 655

RESULT 11

US-10-982-512-19
; Sequence 19, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-19

Query Match 74.6%; Score 3504; DB 5; Length 658;
Best Local Similarity 100.0%; Pred. No. 7.3e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMETEOLGVEIFETADCEENTIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
DB 1 MAAAMETEOLGVEIFETADCEENTIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
QY 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSNGENRNTLFYSEI PKTINRAAVLMLSWKPLL 120
DB 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSNGENRNTLFYSEI PKTINRAAVLMLSWKPLL 120
QY 121 DLFOATLDYGMYSRREELLRRKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSRREELLRRKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVTSCPNIRMDPKLCPADPDWIAFIHSNDIWSNIVTREERLTYVHNEL 240
DB 181 FTQOPLRPNLVTSCPNIRMDPKLCPADPDWIAFIHSNDIWSNIVTREERLTYVHNEL 240
QY 241 ANNEEDARSAGVATFVLOEEDFDRYSGYWCPCAEKTPSPGGKILRIIYEENDESEVEI IHV 300
DB 241 ANNEEDARSAGVATFVLOEEDFDRYSGYWCPCAEKTPSPGGKILRIIYEENDESEVEI IHV 300
QY 301 TSPMLTRADSFYPKTGTPANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
DB 301 TSPMLTRADSFYPKTGTPANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
QY 361 YIARAGWTPEGKYANSLILDRSOTRLQIVLISPELFIPEVDVNMERORLIESVPDSVTPL 420
DB 361 YIARAGWTPEGKYANSLILDRSOTRLQIVLISPELFIPEVDVNMERORLIESVPDSVTPL 420

QY 421 IYVETTDIWINIHDIHVFPQSHHEBIEFIPASECKTGFRHLYKITSILKESKYKRSG 480
DB 421 IYVETTDIWINIHDIHVFPQSHHEBIEFIPASECKTGFRHLYKITSILKESKYKRSG 480
QY 481 GLPAPSDFKCPKKEBIAITSGEWEVLGRHGSNIQVDEVRLRVYFEGTKDSPLEHLYVVS 540
DB 481 GLPAPSDFKCPKKEBIAITSGEWEVLGRHGSNIQVDEVRLRVYFEGTKDSPLEHLYVVS 540
QY 541 YNPGEVTRLDRGYSHSCCISQHCDFPFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
DB 541 YNPGEVTRLDRGYSHSCCISQHCDFPFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
DB 601 WATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655

RESULT 12

US-09-976-674-11
; Sequence 11, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-11

Query Match 74.6%; Score 3504; DB 3; Length 661;
Best Local Similarity 100.0%; Pred. No. 7.3e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSOLKLLADTRKYHYMM 60
DB 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSOLKLLADTRKYHYMM 60
QY 61 AKAPHDFMFVKRNDPGPHSDRIYYLAMSGENRENTLYSEIPKTIINRAAVLMSWKPLL 120
DB 61 AKAPHDFMFVKRNDPGPHSDRIYYLAMSGENRENTLYSEIPKTIINRAAVLMSWKPLL 120
QY 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYVHNEL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYVHNEL 240
QY 241 ANMEEDARSAGVATFVLQEFDRYSGYWCPCAEITTPSGGKILRLIYEENDESEVEIIVH 300
DB 241 ANMEEDARSAGVATFVLQEFDRYSGYWCPCAEITTPSGGKILRLIYEENDESEVEIIVH 300
QY 301 TSPMLETRRADSPRYPKTGTPANKVTFKMSIEMDAEGRIIDVIDKELIQPFELLFEGVE 360
DB 301 TSPMLETRRADSPRYPKTGTPANKVTFKMSIEMDAEGRIIDVIDKELIQPFELLFEGVE 360
QY 361 YIARAGWTPEGKYAWSILLDRSQTRLQIVLISPFLIPVEDDDVMERQRLIESVPDSVTPL 420
DB 361 YIARAGWTPEGKYAWSILLDRSQTRLQIVLISPFLIPVEDDDVMERQRLIESVPDSVTPL 420

QY 421 IYVETTDIWINIHDIHVFPQSHHEBIEFIPASECKTGFRHLYKITSILKESKYKRSG 480
DB 421 IYVETTDIWINIHDIHVFPQSHHEBIEFIPASECKTGFRHLYKITSILKESKYKRSG 480
QY 481 GLPAPSDFKCPKKEBIAITSGEWEVLGRHGSNIQVDEVRLRVYFEGTKDSPLEHLYVVS 540
DB 481 GLPAPSDFKCPKKEBIAITSGEWEVLGRHGSNIQVDEVRLRVYFEGTKDSPLEHLYVVS 540
QY 541 YNPGEVTRLDRGYSHSCCISQHCDFPFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
DB 541 YNPGEVTRLDRGYSHSCCISQHCDFPFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
DB 601 WATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655

RESULT 13

US-10-982-512-11
; Sequence 11, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-11

Query Match 74.6%; Score 3504; DB 5; Length 661;
Best Local Similarity 100.0%; Pred. No. 7.3e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSOLKLLADTRKYHYMM 60
DB 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSOLKLLADTRKYHYMM 60
QY 61 AKAPHDFMFVKRNDPGPHSDRIYYLAMSGENRENTLYSEIPKTIINRAAVLMSWKPLL 120
DB 61 AKAPHDFMFVKRNDPGPHSDRIYYLAMSGENRENTLYSEIPKTIINRAAVLMSWKPLL 120
QY 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYVHNEL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYVHNEL 240
QY 241 ANMEEDARSAGVATFVLQEFDRYSGYWCPCAEITTPSGGKILRLIYEENDESEVEIIVH 300
DB 241 ANMEEDARSAGVATFVLQEFDRYSGYWCPCAEITTPSGGKILRLIYEENDESEVEIIVH 300
QY 301 TSPMLETRRADSPRYPKTGTPANKVTFKMSIEMDAEGRIIDVIDKELIQPFELLFEGVE 360
DB 301 TSPMLETRRADSPRYPKTGTPANKVTFKMSIEMDAEGRIIDVIDKELIQPFELLFEGVE 360
QY 361 YIARAGWTPEGKYAWSILLDRSQTRLQIVLISPFLIPVEDDDVMERQRLIESVPDSVTPL 420
DB 361 YIARAGWTPEGKYAWSILLDRSQTRLQIVLISPFLIPVEDDDVMERQRLIESVPDSVTPL 420

QY 421 IIVEETTDIWINIHDIHVFPQSHEEIEFIFASECKTGRPHLYKITSILKSKYKRSSG 480
Db 421 IIVEETTDIWINIHDIHVFPQSHEEIEFIFASECKTGRPHLYKITSILKSKYKRSSG 480
QY 481 GLPAPSDKCPKEIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS 540
Db 481 GLPAPSDKCPKEIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS 540
QY 541 YNPGEVTRLTDGRGSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db 541 YNPGEVTRLTDGRGSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFYIGG 655

RESULT 14
US-09-976-674-7
; Sequence 7, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-7

Query Match 74.6%; Score 3504; DB 3; Length 690;
Best Local Similarity 100.0%; Pred. No. 7.9e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
Db 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
QY 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLIFYSEI PKTINRAAVLMSWKPLL 120
Db 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLIFYSEI PKTINRAAVLMSWKPLL 120
QY 121 DLFOATLDYGMYSREELLRRERKIGTVGTASDYDHOGSGTFLFOAGSGIYHVKGDPQG 180
Db 121 DLFOATLDYGMYSREELLRRERKIGTVGTASDYDHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLLTYVHNEL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLLTYVHNEL 240
QY 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
Db 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
QY 301 TSPMLETRRADSFYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRRADSFYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSQRLQIVLISPELFPVDDVMERORLIESVPDSVTPL 420
Db 361 YIARAGWTPGKYAWSILLDRSQRLQIVLISPELFPVDDVMERORLIESVPDSVTPL 420

QY 421 IIVEETTDIWINIHDIHVFPQSHEEIEFIFASECKTGRPHLYKITSILKSKYKRSSG 480
Db 421 IIVEETTDIWINIHDIHVFPQSHEEIEFIFASECKTGRPHLYKITSILKSKYKRSSG 480
QY 481 GLPAPSDKCPKEIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS 540
Db 481 GLPAPSDKCPKEIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS 540
QY 541 YNPGEVTRLTDGRGSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db 541 YNPGEVTRLTDGRGSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFYIGG 655

RESULT 15
US-10-982-512-7
; Sequence 7, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-7

Query Match 74.6%; Score 3504; DB 5; Length 690;
Best Local Similarity 100.0%; Pred. No. 7.9e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
Db 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
QY 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLIFYSEI PKTINRAAVLMSWKPLL 120
Db 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLIFYSEI PKTINRAAVLMSWKPLL 120
QY 121 DLFOATLDYGMYSREELLRRERKIGTVGTASDYDHOGSGTFLFOAGSGIYHVKGDPQG 180
Db 121 DLFOATLDYGMYSREELLRRERKIGTVGTASDYDHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLLTYVHNEL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLLTYVHNEL 240
QY 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
Db 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
QY 301 TSPMLETRRADSFYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRRADSFYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSQRLQIVLISPELFPVDDVMERORLIESVPDSVTPL 420
Db 361 YIARAGWTPGKYAWSILLDRSQRLQIVLISPELFPVDDVMERORLIESVPDSVTPL 420

421	IIYBETTDIWINIHDI	FHVPPQSHHEEIEIF	IFASECKTGFRLYKITS	ILKESKYKRSSG	480
421	IIYBETTDIWINIHDI	FHVPPQSHHEEIEIF	IFASECKTGFRLYKITS	ILKESKYKRSSG	480
481	GLPAPSDPKCPI	KEEIAITSGEWEVLGRHGSNTQVDEV	RRLVYFEGTKDPSLEHRLYVVS	540	
481	GLPAPSDPKCPI	KEEIAITSGEWEVLGRHGSNTQVDEV	RRLVYFEGTKDPSLEHRLYVVS	540	
541	YVNPGEVTRLDRGY	SHSICCISQHCDFFI	SKYSNQKNPHCVSLYKLS	SPEDDPTCKTKEF	600
541	YVNPGEVTRLDRGY	SHSICCISQHCDFFI	SKYSNQKNPHCVSLYKLS	SPEDDPTCKTKEF	600
601	WATILDSAGPLPDY	TPPEISFESTTGFTLY	GYMLYKPHDLQPKKYPTVLF	YVGG	655
601	WATILDSAGPLPDY	TPPEISFESTTGFTLY	GYMLYKPHDLQPKKYPTVLF	YVGG	655

Search completed: April 14, 2006, 13:05:34
Job time : 175 secs

SECRET

2024

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 13:02:54 ; Search time 29 Seconds
(without alignments)
1292.259 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMWETQLGVETFTADC.....HLHYLQENLGSRIALKVI 882

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /STDSS/ptodata/2/pubpaa/US08 NEW PUB.pap.*
2: /STDSS/ptodata/2/pubpaa/US06 NEW PUB.pap.*
3: /STDSS/ptodata/2/pubpaa/US07 NEW PUB.pap.*
4: /STDSS/ptodata/2/pubpaa/PCT_NEW PUB.pap.*
5: /STDSS/ptodata/2/pubpaa/US09 NEW PUB.pap.*
6: /STDSS/ptodata/2/pubpaa/US10 NEW PUB.pap.*
7: /STDSS/ptodata/2/pubpaa/US11 NEW PUB.pap.*
8: /STDSS/ptodata/2/pubpaa/US60 NEW PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4700	100.0	882	7	US-11-151-601-20
2	651	13.9	745	7	US-11-079-463-6408
3	577	12.3	109	7	US-11-176-951-10
4	551	11.7	627	7	US-11-079-463-7758
5	530	11.3	762	7	US-11-116-939-13
6	529	11.3	738	7	US-11-208-288-4
7	529	11.3	766	6	US-10-501-035-234
8	529	11.3	766	6	US-11-208-288-2
9	525	11.2	766	6	US-10-522-789-2
10	517.5	11.0	760	7	US-11-208-288-6
11	454	9.7	760	7	US-11-186-284-55
12	394	8.4	99	7	US-11-176-951-11
13	249.5	5.3	657	7	US-11-179-977-1
14	241.5	5.1	624	7	US-11-079-463-7504
15	200.5	4.3	737	7	US-11-079-463-9281
16	145	3.1	102	7	US-11-176-951-7
17	139	3.0	115	7	US-11-176-951-9
18	139	3.0	115	7	US-11-176-951-12
19	133.5	2.8	102	7	US-11-176-951-8
20	125	2.7	24	7	US-11-176-951-16
21	119.5	2.5	1019	6	US-10-995-561-982
22	119	2.5	1307	6	US-10-995-561-711
23	118.5	2.5	710	7	US-11-151-601-23
24	118.5	2.5	1155	7	US-11-098-686-10550
25	118.5	2.5	1243	6	US-10-453-372-1136

Sequence 1134, Ap
Sequence 1142, Ap
Sequence 1132, Ap
Sequence 18, Appl
Sequence 118, App
Sequence 6314, Ap
Sequence 2, Appli
Sequence 143, App
Sequence 11, Appli
Sequence 3, Appli
Sequence 6, Appli
Sequence 14, Appli
Sequence 16, Appli
Sequence 52, Appli
Sequence 709, App
Sequence 4, Appli
Sequence 6, Appli
Sequence 13, Appli

26 118 2.5 1243 6 US-10-453-372-1134
27 118 2.5 4913 6 US-10-453-372-1142
28 118 2.5 4961 6 US-10-453-372-1132
29 115 2.4 849 6 US-10-909-769-18
30 114.5 2.4 668 6 US-10-454-437-118
31 113.5 2.4 420 7 US-11-079-463-6314
32 113.5 2.4 2483 7 US-11-186-999-2
33 112.5 2.4 1285 7 US-11-206-071-2
34 111 2.4 315 7 US-11-018-868-143
35 110.5 2.4 2458 7 US-11-186-999-11
36 110 2.3 1067 7 US/11/062
37 110 2.3 1092 7 US/11/062
38 110 2.3 2455 7 US-11-186-999-14
39 110 2.3 2455 7 US-11-186-999-16
40 109.5 2.3 2725 7 US-11-113-424-52
41 108.5 2.3 1206 6 US-10-995-561-709
42 108.5 2.3 2256 7 US-11-144-368-4
43 108.5 2.3 2455 7 US-11-186-999-4
44 108.5 2.3 2458 7 US-11-186-999-6
45 108.5 2.3 2458 7 US-11-186-999-13

ALIGNMENTS

RESULT 1
US-11-151-601-20
; Sequence 20, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtiss, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MF100-054P1RC10M1D1V1M
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Query Match 100.0%; Score 4700; DB 7; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;

Matches	882;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MAAAMETBQLGVEIFETADCEENIESODRPLEFFVYERYSWSQLKLADTRKRVHGYMM	60						
Db	1	MAAAMETBQLGVEIFETADCEENIESODRPLEFFVYERYSWSQLKLADTRKRVHGYMM	60						
Qy	61	AKAHPDMFVKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTTINRAAVALMSWKPLL	120						
Db	61	AKAHPDMFVKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTTINRAAVALMSWKPLL	120						
Qy	121	DLFOATLDYGMYSREELLERKRGITGVGIASVDYHQSGTFLFQAGSGIYHVKDGPGQ	180						
Db	121	DLFOATLDYGMYSREELLERKRGITGVGIASVDYHQSGTFLFQAGSGIYHVKDGPGQ	180						
Qy	181	FTQOQLRNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWSNIVTREERLTYVHNEL	240						
Db	181	FTQOQLRNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWSNIVTREERLTYVHNEL	240						
Qy	241	ANMEEDARSAGVATFVLOEERDRYSGVWMCPCAKETTPSGGKILRLIYEENDESEVEIIVH	300						
Db	241	ANMEEDARSAGVATFVLOEERDRYSGVWMCPCAKETTPSGGKILRLIYEENDESEVEIIVH	300						
Qy	301	TSPMLETRADSFYPKTGANTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPEILLFEGVE	360						
Db	301	TSPMLETRADSFYPKTGANTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPEILLFEGVE	360						
Qy	361	YIARAGWTPGKYAWSILLDRSQRLQIIVLISPELTFPVEDDVMERORLIESVPSVTPL	420						
Db	361	YIARAGWTPGKYAWSILLDRSQRLQIIVLISPELTFPVEDDVMERORLIESVPSVTPL	420						
Qy	421	IYEBETDWINIHIDIFHVFPQSHEEBIEIFASECKTGPRHLYKITSILKESKYKRSRG	480						
Db	421	IYEBETDWINIHIDIFHVFPQSHEEBIEIFASECKTGPRHLYKITSILKESKYKRSRG	480						
Qy	481	GLPAPSDFKCPIKEBIATTSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHHLIYVVS	540						
Db	481	GLPAPSDFKCPIKEBIATTSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHHLIYVVS	540						
Qy	541	YVNPGEVRLTRDRGYSHSCCISQHCDFFIISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF	600						
Db	541	YVNPGEVRLTRDRGYSHSCCISQHCDFFIISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF	600						
Qy	601	WATILDSAGPLPDYTPPPIFSESTTGFTLYGMLYKPHDLQPGKKYPTVLFYGGPQVOL	660						
Db	601	WATILDSAGPLPDYTPPPIFSESTTGFTLYGMLYKPHDLQPGKKYPTVLFYGGPQVOL	660						
Qy	661	VNNRFGVKYPRNLTLASLGYVVVVVVDNRGSGCHRGKLFEGAFKYMKGQIEIDDOVEGLQY	720						
Db	661	VNNRFGVKYPRNLTLASLGYVVVVVVDNRGSGCHRGKLFEGAFKYMKGQIEIDDOVEGLQY	720						
Qy	721	LASRYDFIDLDVRVGIHGSYGGYLSLWALMORSIDIPRAIAGAPVTLWIFDYGTYERYM	780						
Db	721	LASRYDFIDLDVRVGIHGSYGGYLSLWALMORSIDIPRAIAGAPVTLWIFDYGTYERYM	780						
Qy	781	GHPDQNEGYVILGSAVMAOAEKFPSEPNLHLLHGLFDENVHFAHTSILLSFLVRAGKPYD	840						
Db	781	GHPDQNEGYVILGSAVMAOAEKFPSEPNLHLLHGLFDENVHFAHTSILLSFLVRAGKPYD	840						
Qy	841	LQIYPOERHSIRVPSGEGHYELHLLHYLOENLGSRIAAKVI	882						
Db	841	LQIYPOERHSIRVPSGEGHYELHLLHYLOENLGSRIAAKVI	882						

RESULT 2
US-11-079-463-6408
; Sequence 6408, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463

; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6408
; LENGTH: 745
; TYPE: PRT
; ORGANISM: B.fragilis
; US-11-079-463-6408

Query Match 13.9%; Score 651; DB 7; Length 745;
Best Local Similarity 27.3%; Pred. No. 2e-47;
Matches 199; Conservative 124; Mismatches 244; Indels 162; Gaps 29;

Qy	173	VKDGPGQFTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWSNIVTREER	232
Db	150	LSDGGPQ-----QAPVFS-DGNLVAFVRDNNFLV-----K	180
Qy	233	LTVVNELANMEEDAR----SAGVATFVLOE--FDRYSGVWMCPCAKETTPSGGKILRL	286
Db	181	LLYGNSE-SQVTEDEKLSVLNGIPDWYBEEFGFNRALEF----NADNT-----MLA	228
Qy	287	YEENDESEVEIIVHVSFPLM--ETRRAD-----SFYRYPKTGTANPKV---TF-----	327
Db	229	YVRFDESEVP--SYTFPLFAGEAPRYDALQDYPGEYTYKYPKAGYPNKSUVHTFDIKSK	286
Qy	328	--KMSEIMIDAREGIIDVIDKELIQPEILLFEGVEIARAGWTPGKYAWSILLDRSQTR	385
Db	287	VTRQVKLPIDAG-----YIPRIRFTQDPNKLAIMTLNRHQR	324
Qy	386	LQIVLISPELTFPVEDDVMERORLIESVPSVTPLIIVEETDIIWNIHIDIFHVFPQSHE	445
Db	325	FMVYFADP-----RSTVCKLALRDESFPYINE-----NVFDNIQFYPE---362	
Qy	446	BEIEFIPASECKTGPRHLYKITSILKESKYKRSRGGLPAPSDFKCPIKEBIATTSGEWEV	505
Db	363	---YFSFVSD-KSGIPHLIY-----WYSMNGNL-----IKQ---VTSNGYEV	396
Qy	506	LGHGNSIQVDEVRLVYFEGTKDSPLEHHLIYVSYVNPGEVRLTRDRGYSHSCCISQHC	565
Db	397	KNFIGWNPDTNE----FYTSNEESFMRQAVIKID--RKGKMKLSQPGTNSPIFSSSM	450
Qy	566	DFPISKYSNQKNPHCVSLYKLSPEDDPTCKTKEFWATILDSAGPLPDYTPP--EIFSFE	623
Db	451	KYPMNKFTSLDTPMLITL-----NDNTGKVLKTLVTNDKCLKOKLAEBYAIPOKEFFTK	503
Qy	624	STTGFTLYGMLYKPHDLQPGKKYPTVLFYGGPQVOLVNNRFGVKYPRNLTLASLGYV	683
Db	504	TTEGVDLNGMMMKPVNFPDAKYPVLMFQYSGSGSQVLDKW-GISW--ETYMASLGYV	560
Qy	684	VVIDNRGSGCHRGKLFEGAFKYMKGQIEIDDOVEGLQYLASRYDFIDLDVRVGIHGSYGGY	743
Db	561	ACVDGRGTGGRGSEFQKCTYLNGLVKEARQDVEAAKYLGG-LPYVDKGRIGIWMGSGGY	619
Qy	744	LSLMALMORSIDIPRAIAGAPVTLWIFDYGTYERYMGHPDQNEGYVILGSAVMAOAEKFP	803
Db	620	MTIMSMSEGTVPFKAGVAAPTDWKYDVTYTERFMRTPKENAEGYKAASAFSAADNL-	678
Qy	804	SEPNRLHLLHGLFDENVHFAHTSILLSFLVRAGKPYDQIYPOERHSIRVPSGEGHYELH	863
Db	679	--HGNLLVHGMADDNVHVFQNCTEVAEHLVOLGQFDMQVTVNRRNHSIYGNGNTRHLYTK	736
Qy	864	LLHYLOENL	872
Db	737	LTNFFRNLL	745

RESULT 3
US-11-176-951-10
; Sequence 10, Application US/11176951
; Publication No. US20060024313A1

GENERAL INFORMATION:
APPLICANT: CHEN, XIN
TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
PROLYD PEPTIDASES
FILE REFERENCE: 08842.0019
CURRENT APPLICATION NUMBER: US/11/176,951
CURRENT FILING DATE: 2005-07-06
PRIOR APPLICATION NUMBER: 60/586,095
PRIOR FILING DATE: 2004-07-06
PRIOR APPLICATION NUMBER: 60/585,952
PRIOR FILING DATE: 2004-07-06
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patent in Ver. 3.3
SEQ ID NO 10
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapiens
US-11-176-951-10

Query Match 12.3%; Score 577; DB 7; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.5e-42;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 774 GYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRLLHLLHGFLENVHFAHTSILLSPLV 833
DB 1 GYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRLLHLLHGFLENVHFAHTSILLSPLV 60
QY 834 RAGKPYDQIYPOERHSIRVPSGEHYELHLLHLYQNLGSRGAALKVI 882
DB 61 RAGKPYDQIYPOERHSIRVPSGEHYELHLLHLYQNLGSRGAALKVI 109

RESULT 4
US-11-079-463-7758
Sequence 7758, Application US/11079463
Publication No. US20060073161A1
GENERAL INFORMATION:
APPLICANT: Gary L. Breton
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
PATHOGENICITY
FILE REFERENCE: PATH00-03DIV2
CURRENT APPLICATION NUMBER: US/11/079,463
CURRENT FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/128,705
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: US 09/540,209
PRIOR FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 10444
SEQ ID NO 7758
LENGTH: 627
TYPE: PRT
ORGANISM: B. fragilis
US-11-079-463-7758

Query Match 11.7%; Score 551; DB 7; Length 627;
Best Local Similarity 25.7%; Pred. No. 6.3e-39;
Matches 168; Conservative 120; Mismatches 249; Indels 116; Gaps 23;
QY 235 YVHNLNANMEEDARSAGVATFVLOQEFDRYSYVWCPKAETTPSGGKILRILEYENDESE 294
DB 76 YVDDRAVNEPGEIVCGQS--VHRNEFGIKKGTW-----SPSGNLLAFYRMDQSWAQ 127
QY 295 VRIHVTSPMLRETRADSRYPKGTGTANPKVTFKMSIMDAEGRII-----DVIDKELI 349
DB 128 YPLVDVTAIAE--VNNIRYPMAGMTSHQKV--GIYNPATKSIYLNAGDPTDR--- 178
QY 350 QPFEILFEGVEYIARAGWTPEGKAVMSILLRSQTRQLVLSPELFIPEVDDVMERQRL 409
DB 179 -----YFTNISWAPDEKSLYLIELNRDQNHAKLCRYD----- 210
QY 410 IESVPSVPTLLIYEETDWINIHDIHFVPSQHEEIEIPFASECKTGFRHLYKITSI 469

DB 211 VETGELTAT---LFEEKSDKYVEPODPIIFLPWONS---KFIYOSQ-KQGFSHLYLYDTN 263
QY 470 LKESKYRSGGLPAPSPDFKCPKEIEAITSGEW---EVLGRHGSNIQDEVRLRVFEG 526
DB 264 GROIR-----OLTEGDMIVKEVLG-----FDTKKEIIAS 294
QY 527 TKDSPLEHLYVVSYNVPEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKL 586
DB 295 TEFSPLONNLFRLD-TKTGTRTFLGASGVHSGQLSPGRYLIDQYNSPTVPSRINIIV 353
QY 587 SSPEDDPTCKTEFWATILDSAGPLPDYTPP--EIFSESTTGT-LYGMLYKPHDLQPG 643
DB 354 QSGKS-----VNLTLAADPTGYKMPGIEGTIKAADGKTIDLYRLIKPADFDPN 403
QY 644 KKYPTVLFIYGPQVQLVNNREF-GVKYFRNLNTLASLGVVVVVVDNDSCHRGKLPFGAF 702
DB 404 KKYPAIVVYVGGPHALVTNGQNGARGWDI-YMANKGYIMFTVDGSSNRGLDFENVT 462
QY 703 KYKMGQIEIDDOVGLQVLAISRYDFIDLDRIYHGSYGGYLSLMAWMQSDIFRVAIAG 762
DB 463 FROLGIEGRDQVKGTEFLKS-LPYVDGNRIGVHGSFGGHTTALLRYPEIFKVGAG 521
QY 763 APVTLMIFVDTGTYERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRLLHLLHGFLENVH 822
DB 522 GPVIDGYYEVMYGERYMDTPQSNPKGYKECNLKNLGNL---KGHLMIHDDHDDTVP 578
QY 823 AHTSILLSFLVRAKPYDQIYPOERHSIRVPSGE---HYELHLLHLYQNL 872
DB 579 QHTLSFMKACIDARTYVDLFIYPCHKNV---SGDRVHLHEKITRYFEDYL 627

RESULT 5
US-11-116-939-13
Sequence 13, Application US/11116939
Publication No. US20050265995A1
GENERAL INFORMATION:
APPLICANT: Stephen Tomlinson
APPLICANT: Richard J. Quigg
TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
FILE REFERENCE: 19113.0115U2
CURRENT APPLICATION NUMBER: US/11/116,939
CURRENT FILING DATE: 2005-04-28
PRIOR APPLICATION NUMBER: 60/565,907
PRIOR FILING DATE: 2004-04-28
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 762
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
OTHER INFORMATION: construct
US-11-116-939-13

Query Match 11.3%; Score 530; DB 7; Length 762;
Best Local Similarity 24.0%; Pred. No. 5.5e-37;
Matches 218; Conservative 132; Mismatches 318; Indels 240; Gaps 45;
QY 17 TADCEENIESQDRPKLEPFYFVERYSWSQLKL--LADTRKYHYMMKAPAHDFMFKEND 74
DB 36 TADSRKTYTLTD-----YLNKY-----LKLRLWISD-----HEYLYKOENN 71
QY 75 PDGPHSDRIYLLAWSGENRENTLF-----YSRIPKTIINRAAVLMSWPKLLDLFOATLDY 129
DB 72 -----ILVFNAYGNSSVLENSTFDFGHSIN-----DY 101
QY 130 GMYSRBEELLRKRKIGTVGIASDY-----HOGSGTFLFOAGSGIYHVKGQPGGTQQ 184
DB 102 SISPDGOFILLE-----YNYVKQWRHS-----YTASYDIYDLNK----- 135
QY 185 PLRPNLVETSCPNIRMDPKLCPADPDWI-----AFIHSNDIWINIVTRERRRLTY 235

Db 136 --RQLITERIPN-----NTQWTVSPVGHKLAYVWNNDIYVKIEPNLPSYRITW 183
Qy 236 VHNELANNEEDARAGATVFLQEE-FDRYSGYVWGPCAETTPSGGKIILRIIYENDESE 294
Db 184 TG-----KEDIINYITDWTVEEVSFASALWNSFNGTF-----LAYAFNDTE 228
Qy 295 VEIIH---VTSFMLETRADSFYPTKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQP 351
Db 229 VPLIBSYSDESIQPKTVKVPYKAGAVNPTKFFV--VNTDSLSSVTNATSQITAP 286
Qy 352 FEILPEGVEYIARAGWTPEGKYAWSIILLDRSOTRLQIVLISPELFIPEVDDVMERQRIE 411
Db 287 ASMLI-GHYLCDVTWA-----TOERISLOWL-----RRIQ 316
Qy 412 SVPDQSVTLIIYEETDIWINIHDFHVPFQSHHEEIEFIFASECKTGFRH-----LY 464
Db 317 NY--SVMDICDYDESSGRW-NC-----LVARQHIEMSTTGWGRPRPSPHFTLDGNSFY 368
Qy 465 KITSILKESKYKRSSGGLPAPSPDFKCPIKEEIAITSGEWELGRHGSNIQVDEVRRLVVF 524
Db 369 KIIS--NEEGYRHI-----CYFQIDKDKCTFITKGTWEVIG-----LEALTSYLYYI 414
Qy 525 EGT-KDSPLEHHLVYVSVNPGEVTRLTDRGYSHSCIS-----OHCDFPISKYSNOKNP 578
Db 415 SNEYKMGCGRNLYKI-----QLSD--YTKVTLSCELNPERCQIYVSFSKEAKY 463
Qy 579 H---C-----VSLYKLSPPEDPTCKTEFWATILDSAGPLPDYTPPE-----IFSPESTGCF 628
Db 464 YQRCGSGPLPLYTLHSSVNDKGLRVLED-NSALDKM--LQNVQMFSSKLDIILNETKF 520
Qy 629 TLYGMLYKPHLDQPKKYPTVLFITYGQPVQLVNNRFPKGVKYFRLN---TLASLGVVVV 685
Db 521 -WYQMLPPH-FDKSKYKPLLDVYAGPCSQ-----KADTVFLNWAATYLASTENIIA 572
Qy 686 -IDNRGSHRGLKFGAFKYKMGQIEIDQVEGLQYLASRYDFIDLDVRGIIHWSYGGYL 744
Db 573 SPDGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQF-SKMGFVDNKRITAIWGSYGGV 631
Qy 745 SLMALMQRSDFIRVAIAGAPVTLWIFDYDTGTYERYMG--HPDQNEQGYLGSVAMQAKF 802
Db 632 TSMVLGSGGVFKCGIAVAPVSRWEYDVSVTERYMGVGLPTPEDNLDHYRNSTVMSRAENF 691
Qy 803 PSEPNRLLHGHFDENVHFHAHTSILLSFLVRAGKPYDLQIYPOERHSIRVPSGEGHYEL 862
Db 692 --KQVEYLLIHGTADDNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAQHITYT 749
Qy 863 HLHLYLOE 870
Db 750 HNSHFQIK 757

RESULT 6

US-11-208-288-4
; Sequence 4, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-4
Query Match 11.3%; Score 529; DB 7; Length 738;

Best Local Similarity 26.5%; Pred. No. 6.3e-37;
Matches 186; Conservative 109; Mismatches 200; Indels 128; Gaps 32;
Qy 206 PADPQWI-----AFIHSNDIWSINIVTREERRLTIVHNELANNEEDARSGVATFV 256
Db 121 PNTQWTVSPVGHKLAYVWNNDIYVKIEPNLPSYRITWTG-----KEDIINYITDWTW 174
Qy 257 LOEE-PDRYSGYVWGPCAETTPSGGKIILRIIYENDESEVEIIH---VTSFMLETRADS 312
Db 175 YEEVEVSFASALWNSFNGTF-----LAYAFNDTEVPLIEBYSYSDESIQPKTVR 225
Qy 313 FRYPTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPEILPEGVEYIARAGWTPEGK 372
Db 226 VYPKAGAVNPTKFFV--VNTDSLSSVTNATSQITAPASMLI-GHYLCDVTWA--- 278
Qy 373 YAWSIILLDRSOTRLQIVLISPELFIPEVDDVMERQRIESVPSDVTPLIIYEETDIWIN 432
Db 279 -----TOERISLOWL-----RRIQNY--SVMDICDYDESSGRW-N 310
Qy 433 IHDIFHVPFQSHHEEIEFIFASECKTGFRH-----LYKITSILKESKYKRSSGGLPAP 485
Db 311 C-----LVARQHIEMSTTGWGRPRPSPHFTLDGNSFYKIIS--NEEGYRHI----- 356
Qy 486 SDPKCPIKEEIAITSGEWELGRHGSNIQVDEVRRLVVFECT-KDSPLEHHLVYVSVNPN 544
Db 357 CYFQIDKDKCTFITKGTWEVIG-----LEALTSYLYYISNEYKMGCGRNLYKIOLIDY 411
Qy 545 GEVTRLTDRGYSHSCIS-OHCDFPISKYSNOKNPH---C-----VSLYKLSPPEDDPTCK 596
Db 412 TKVTCL-----SCELNPERCQIYVSFSKEAKYQLRCSGPGCLPLYTLHSSVNDKGLR 464
Qy 597 TKEFWATILDSAGPLPDYTPPE---IFSPESTGFTLYGMLYKPHLDQPKKYPTVLFYI 653
Db 465 VLED-NSALDKM--LQNVQMFSSKLDIILNETKF-WYQMLPPH-FDKSKYKPLLDVY 519
Qy 654 GGPQVQLVNNRFPKGVKYFRLN---TLASLGVVVV-INDRGSHRGLKFGAFKYKMGQI 709
Db 520 AGPCSQ-----KADTVFLNWAATYLASTENIIIVASFDGRSGYQGDKIMHAINRRLGT 573
Qy 710 EIDDQVEGLQYLASRYDFIDLDVRGIIHWSYGGYLSMALMQRSDFIRVAIAGAPVTLWI 769
Db 574 EVEDQIEAARQF-SKMGFVDNKRITAIWGSYGGVTSVNLGSGGVFKCGIAVAPVSRWE 632
Qy 770 FDYDTGTYERYMG--HPDQNEQGYLGSVAMQAKFPPSEPNRLLHGHFDENVHFHAHTSI 827
Db 633 YDVSVTERYMGVGLPTPEDNLDHYRNSTVMSRAENF--KQVEYLLIHGTADDNVHFQSSAQ 690
Qy 828 LLSFLVRAGKPYDLQIYPOERHSIRVPSGEGHYELHLHLYLOE 870
Db 691 ISKALVDVGVDFQAMWYTDDEHGIASSTAQHITYHNSHFQIK 733

RESULT 7

US-10-501-035-234
; Sequence 234, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 234
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-501-035-234

Query Match 11.3%; Score 529; DB 6; Length 766;
Best Local Similarity 26.5%; Pred. No. 6.7e-37;
Matches 186; Conservative 109; Mismatches 280; Indels 128; Gaps 32;

QY 206 PADPDWI-----AFHSNDIWSNIVTTRERRLLTYVHNLANMEEDARSAGVATFV 256
DB 149 PNTQWVWSPVGHKLAYWNNDIYVKIEPNLPSYRIWTG-----KEDIYNGITDMV 202
QY 257 LQEE-FDRYGYWVWCPKAEPTTSGGKILRLIYEENDESEVEIHH---VTSPLMETRRADS 312
DB 203 YEEVEFSAYSALWWSNGTF-----LAVAQNDTEVPLIEYSFVSDESLOYPKTVR 253
QY 313 FRYPKGTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIOPFELLFEGVEYIARAGWTPGK 372
DB 254 VPYPKAGAVNPTVKFV--VNTDSLSSVTNATSIQITAPASMLI-GDHYLCDVTWA----- 306
QY 373 YAWSILLDRSQRLQIOLVILSPFLFIPVEDDVWERQRLIESVPDSVTPLIIYEETDIIWIN 432
DB 307 -----TQERISLQWL-----RRIQNY--SYMDICDYDESSGRW-N 338
QY 433 IHDIFHVFPQSHEEIEFIFASECKTGPRH-----LYKITSILKESKYKRSSGGLPAP 485
DB 339 C-----LVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIIS--NEEGYRHI----- 384
QY 486 SDFKCPKEIAITSGEWELGRHGSNIQVDEVRRLVYFEGT-KDSPLEHHLVYVSVNP 544
DB 385 CYFOIDKKOCTFITKGTWVIG-----IEALTSYLYIISNEYKMGPGGRNLYKIQLIDY 439
QY 545 GEVTRLDRGYSHSCCIS--QHCDFFISKYSNQKNPH---C-----VSLYKLSPPDDPTCK 596
DB 440 TKVTCL-----SCELNPERCOYYSVSFEAKYQRLCSGPGPLPLYTLHSSVNDKGLR 492
QY 597 TKFWMATILDSAGPLPDYTPPE---IFSPESTGTGLYGMLYKPHDLQPKKYPTVLFIY 653
DB 493 VLED-NSALDKM--LQNVQWPSKKLDFIILNETKF-WYQMLPPH-FDKSKKYPILLDVIY 547
QY 654 GGPQVQLVNNRFGVKYKFLN---TLASLGYYVVV-IDNRGSHRGKLFEGAFKYMGOI 709
DB 548 AGPCSQ-----KADTVFLRNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRLGTF 601
QY 710 EIDDOVEGLQYLASRYDFDLDVRGIGHWSYGYLSLMAQMORSDFRVAIAGAPVTLWI 769
DB 602 EVEDQIEAARQF-SKMGFVNDKRIATWGSYGYVTSVWLGSVGFVKCGIAPVSRWE 660
QY 770 FYDTGYTERYMG--HPDQNEQGYLGSVAMQAEKPPSPNRLLLHGFLDENVHFAHTSI 827
DB 661 YDSVYTERYMGUPTPEDNLDHYNSTVMSRAENF--KQVEYLLHGTADDNVHFOQSAQ 718
QY 828 LLSFLVRAGKPYDLQIYPOERHSIRVPESGEHYELHLLHYLQE 870
DB 719 ISKALVDVGVDFQAMWYTDHGHASSTAHQHIYTHMSHFQ 761

RESULT 8

US-11-208-288-2

; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens

US-11-208-288-2

Query Match 11.3%; Score 529; DB 7; Length 766;
Best Local Similarity 26.5%; Pred. No. 6.7e-37;
Matches 186; Conservative 109; Mismatches 280; Indels 128; Gaps 32;

QY 206 PADPDWI-----AFHSNDIWSNIVTTRERRLLTYVHNLANMEEDARSAGVATFV 256
DB 149 PNTQWVWSPVGHKLAYWNNDIYVKIEPNLPSYRIWTG-----KEDIYNGITDMV 202
QY 257 LQEE-FDRYGYWVWCPKAEPTTSGGKILRLIYEENDESEVEIHH---VTSPLMETRRADS 312
DB 203 YEEVEFSAYSALWWSNGTF-----LAVAQNDTEVPLIEYSFVSDESLOYPKTVR 253
QY 313 FRYPKGTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIOPFELLFEGVEYIARAGWTPGK 372
DB 254 VPYPKAGAVNPTVKFV--VNTDSLSSVTNATSIQITAPASMLI-GDHYLCDVTWA----- 306
QY 373 YAWSILLDRSQRLQIOLVILSPFLFIPVEDDVWERQRLIESVPDSVTPLIIYEETDIIWIN 432
DB 307 -----TQERISLQWL-----RRIQNY--SYMDICDYDESSGRW-N 338
QY 433 IHDIFHVFPQSHEEIEFIFASECKTGPRH-----LYKITSILKESKYKRSSGGLPAP 485
DB 339 C-----LVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIIS--NEEGYRHI----- 384
QY 486 SDFKCPKEIAITSGEWELGRHGSNIQVDEVRRLVYFEGT-KDSPLEHHLVYVSVNP 544
DB 385 CYFOIDKKOCTFITKGTWVIG-----IEALTSYLYIISNEYKMGPGGRNLYKIQLIDY 439
QY 545 GEVTRLDRGYSHSCCIS--QHCDFFISKYSNQKNPH---C-----VSLYKLSPPDDPTCK 596
DB 440 TKVTCL-----SCELNPERCOYYSVSFEAKYQRLCSGPGPLPLYTLHSSVNDKGLR 492
QY 597 TKFWMATILDSAGPLPDYTPPE---IFSPESTGTGLYGMLYKPHDLQPKKYPTVLFIY 653
DB 493 VLED-NSALDKM--LQNVQWPSKKLDFIILNETKF-WYQMLPPH-FDKSKKYPILLDVIY 547
QY 654 GGPQVQLVNNRFGVKYKFLN---TLASLGYYVVV-IDNRGSHRGKLFEGAFKYMGOI 709
DB 548 AGPCSQ-----KADTVFLRNWATYLASTENIIVASFDGRGSGYQGDKIMHAINRRLGTF 601
QY 710 EIDDOVEGLQYLASRYDFDLDVRGIGHWSYGYLSLMAQMORSDFRVAIAGAPVTLWI 769
DB 602 EVEDQIEAARQF-SKMGFVNDKRIATWGSYGYVTSVWLGSVGFVKCGIAPVSRWE 660
QY 770 FYDTGYTERYMG--HPDQNEQGYLGSVAMQAEKPPSPNRLLLHGFLDENVHFAHTSI 827
DB 661 YDSVYTERYMGUPTPEDNLDHYNSTVMSRAENF--KQVEYLLHGTADDNVHFOQSAQ 718
QY 828 LLSFLVRAGKPYDLQIYPOERHSIRVPESGEHYELHLLHYLQE 870
DB 719 ISKALVDVGVDFQAMWYTDHGHASSTAHQHIYTHMSHFQ 761

RESULT 9

US-10-522-789-2

; Sequence 2, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIVAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens

APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Thibodeau, Stephen N.
APPLICANT: Burgat, Lawrence J.
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF COLON CANCER
FILE REFERENCE: MP01-029P2RNM
CURRENT APPLICATION NUMBER: US/11/186,284
PRIOR FILING DATE: 2005-07-21
PRIOR APPLICATION NUMBER: US/10/301,822
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/339,971
PRIOR FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: US 60/361,978
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/381,988
PRIOR FILING DATE: 2002-05-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 55
LENGTH: 760
TYPE: PRT
ORGANISM: Homo Sapiens
US-11-186-284-55

Query Match 9.7%; Score 454; DB 7; Length 760;
Best Local Similarity 24.1%; Pred. No. 1.9e-30; Indels 150; Gaps 27;
Matches 169; Conservative 113; Mismatches 269;
QY 212 IAFHSNDIWSNITVREERRLYVHNELANMEEDARSAGVATFVLQEEF--DRYSGYWM 269
DB 162 LAYVQNNIYLKQRPQPPQITF-----NGRENKIFNGIPDWVVEEMLPKYA-LWM 214
QY 270 CPKAETPSGKILIRLYENDESEVEIHHVTSPMLET-RRADSFYKPTGTANTKVFTEK 328
DB 215 SP-----NGKFL--AYAEFNDKDIPIAYSYVGDQYPTINIPYKAGAKNFVARI- 264
QY 329 MSEIMDAEGLIIVDKELQPEI-----LPEGVETARAGWTPEGKYAVSILLDRSQ 383
DB 265 -----FIIDTTPYAVGVQEPVVPVPMIASSDYFSLWTVTWDERVCLQW----- 309
QY 384 TRLQIVLISPELFIPEVDVMMERQRLIESVPDSVT-----PLIYBETTDIWINI 433
DB 310 KVVQNVSVLSCDFREDQWDCPKTOEHIEESRTGWAGGFVSRPVFSYDA-----ISY 364
QY 434 HDIFHPVQSHHEIEFIFASECKTGFRHLYKITSILKESKYRSGGLPAPSDPKCPK 493
DB 365 VKIF-----SDKDGKHIHYI-----KDTVE 385
QY 494 BEIATSGEWEVLGRHGSNIQVDEVRRLVVEFGTKDPSLEHLYVVSYPNPGVETRLTDR 553
DB 386 NAIQITSGKWEAI-----NI-----FRVQDSLFYSNFEPEYPCRRNIVRISIG 430
QY 554 GYSHS-CCISOH-----CDFRISKYSNKNPHCVSLY-----KLSSPEDDPTCKTFEWMAT 603
DB 431 SYPPSKCVTHLKRQCOYTFASDYAKYALVCYGPICPISLHDGRTDQ-----IK 486
QY 604 ILDSAGPLD-----YTPPEIFSPSTGFTLYGMVLYKPHDLQPGKKYPTVLFYGGPOV 658
DB 487 ILEENKELNALKNQLPKEIKKLEVDEITLWYKMLPQPQDRSKKYPILLIQVGGPCS 546
QY 659 QLVNRFKGVKYLRLNTLASL-----GYVVVIDNRGSGHCHLPEGAFKYGQIETDQ 714
DB 547 QSVRS-----VFAVNWISYLSKSGVMVIALVDGRGTAFQDGLLVAVRKLGUYVEVDQ 600
QY 715 VEGLOYLASRYDFIDLDVGRHGSYGGYLSLMAIMQSDIFRVAIAGAPVTLWIFYDTG 774
DB 601 ITAVRKF1-EMGFIDEXIALWHSYGGYSSIALASGTGLFKCGIAPVSSWEYASV 659
QY 775 YTERYMGHP--DQNEQGYLGSVAMQAKFPSEPNRLLLHGLFDENVHFAHTSILLSP 832
DB 660 YTERFMGLPTKDDNLHKNSTVMARAEYFRNVD--YLLIHGTADDNVHVFQNSAQIAKAL 717

QY 833 VRAGKPYDLQIYQERHSIRVPESG---EHVELLHLHYLOE 870
DB 718 VNAQVDFQAMWYSDQHGL-----SGLSTNHLTYHTMTFLKQ 754
RESULT 12
US-11-176-951-11
Sequence 11, Application US/11/176951
Publication No. US20060024313A1
GENERAL INFORMATION:
APPLICANT: CHEN, XIN
APPLICANT: CHEN, YUAN-SHOU
TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
PROLYL DIPEPTIDASES
FILE REFERENCE: 08842.0019
CURRENT APPLICATION NUMBER: US/11/176,951
CURRENT FILING DATE: 2005-07-06
PRIOR APPLICATION NUMBER: 60/586,095
PRIOR FILING DATE: 2004-07-06
PRIOR APPLICATION NUMBER: 60/585,952
PRIOR FILING DATE: 2004-07-06
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 11
LENGTH: 99
TYPE: PRT
ORGANISM: Homo sapiens
US-11-176-951-11

Query Match 8.4%; Score 394; DB 7; Length 99;
Best Local Similarity 72.7%; Pred. No. 1.2e-26;
Matches 72; Conservative 10; Mismatches 17; Indels 0; Gaps 0;
QY 774 GYTERYMGHPDQNEGYLYGSVAMQAKFPSEPNRLLLHGLFDENVHFAHTSILLSP 833
DB 1 GYTERYMGHPDQNEGYLYGSVAMQAKFPSEPNRLLLHGLFDENVHFAHTSILLSP 60
QY 834 RAGKPYDLQIYQERHSIRVPESGHEHYELHLLHYLOENL 872
DB 61 RAGKPYDLQIYQERHSIRVPESGHEHYELHLLHYLOENL 99

RESULT 13
US-11-179-977-1
Sequence 1, Application US/11/179977
Publication No. US20050249789A1
GENERAL INFORMATION:
APPLICANT: Genencor International, Inc.
TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
FILE REFERENCE: GC511-PCT
CURRENT APPLICATION NUMBER: US/11/179,977
CURRENT FILING DATE: 2005-07-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 657
TYPE: PRT
ORGANISM: Bacillus
US-11-179-977-1

Query Match 5.3%; Score 249.5; DB 7; Length 657;
Best Local Similarity 21.4%; Pred. No. 6e-13;
Matches 150; Conservative 101; Mismatches 265; Indels 185; Gaps 31;
QY 245 EDARSAGVATFVLQEEFDRYSGYVWCPKAEI-----TPSGGKILRL 286
DB 24 DGTAAVYKSVQVQNEK-DSYTNIMYETKGTGGSVPWTHGKRSRSTDPKWSPDGRTLAFIS 82
QY 287 YEENDESEVEIHHVTSPMLETRADSFYKPTGTANTKVFQKSEIMID---ABGRITDV 343
DB 83 DRGDAQAQLYIMSTEGE--EARKLTDIPY---GVSKPLWSPDGESILVTSLGESIDD. 137

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 00:12:06 ; Search time 11157 Seconds

(without alignments)
4493.674 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAAMETFEQLGVEIFETADC.....HLLHYLQENLGRIRALXVI 882

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abses/ABSSWEB_spool/US10825632/runat_14042006_092559_4657/app_query.fasta.1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abses08
-USER=US10825632 @CGN 1 1 4939 @runat_14042006_092559_4657 -NCFU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	4700	100.0	2649	6 AX354795	AX354795 Sequence
2	4700	100.0	2656	8 AV172659	AV172659 Homo sapi
3	4700	100.0	2671	6 AR631278	AR631278 Sequence

ALIGNMENTS

RESULT 1	AX354795	Sequence 3 from Patent WO0179473.	2649 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX354795	Sequence 3 from Patent WO0179473.	2649 bp	DNA	linear	PAT 06-FEB-2002
DEFINITION	AX354795	Sequence 3 from Patent WO0179473.	2649 bp	DNA	linear	PAT 06-FEB-2002
ACCESSION	AX354795	Sequence 3 from Patent WO0179473.	2649 bp	DNA	linear	PAT 06-FEB-2002
VERSION	AX354795.1	GI:18619528	2649 bp	DNA	linear	PAT 06-FEB-2002
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1	Meyers, R.A. and Williamson, M.				
AUTHORS	21953,	a human prolyl oligopeptidase family member and uses thereof				
TITLE	Patent: WO 0179473-A 3	25-OCT-2001;				
JOURNAL	Millennium Pharmaceuticals, Inc. (US)					
FEATURES	Location/Qualifiers					
source	1. .2649	/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				

ORIGIN	Alignment Scores:	0	Length:	2649
	Pred. No.:	4700.00	Matches:	882
	Score:			

Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0
US-10-825-632-1 (1-882) x AX354795 (1-2649)			
Qy	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	1	ATGCACACACATGGAAACAGAACAGCTGGGTGGATATTTGAAATTCGGGACTGT	60
Qy	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluuArgTyr	40
Db	61	GAGGAGAAATTTGAATCAGAGATCGGCTTAATTTGGAGCCCTTTTATGTTGACGGTAT	120
Qy	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Db	121	TCCTGGAGTCAGCTTTAAAGCTGCTTGCCTGCGATACAGAAATATCATCGCTACATGATG	180
Qy	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Db	181	GCTAAGCACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATITCA	240
Qy	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Db	241	GACAGAATCTATTACCTTGGCATGCTCTGGTGAACAGAGAAAATACACTGTTTATTCT	300
Qy	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Db	301	GAATTTCCCAAAATATCAATAGACGACGAGTCTTAATGCTCTCTTGGAAAGCCCTTTTGG	360
Qy	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Db	361	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAAGAAAGAACTATTAGA	420
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Db	421	GAAGAAACCGCATTTGGACAGTCGGAAATTTGCTTCTTACGATTATCAACCAAGAGTGA	480
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Db	481	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGGA	540
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Db	541	TTTACGCCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTTCCCAACATACGGATG	600
Qy	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Db	601	GATCCAAATATTATGCCCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATAT	660
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisnGluLeu	240
Db	661	TGGATATCTTAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTA	720
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Db	721	GCCAAACATGGAAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA	780
Qy	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Db	781	TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCAAGTGTGTGT	840
Qy	281	LysIleLeuArgIleLeuTyrGluLysAsnAspGluSerGluValGluIleIleHisVal	300
Db	841	AAAAATCTTAGAATTTCTATATGAAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT	900
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Db	901	ACATCCCTCATGTTGGAAACAGAGGGGACAGATTCATTCCTGTTATCTTAAACAGGTACA	960
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Db			
Db	961	GCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGAAGGAGGATC	1020
Qy	341	IleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Db	1021	ATGATGTCATAGATAAGGAACATAATTTCAACCTTTTGAAGATTCATTTTGAAGAGTTGAA	1080
Qy	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
Db	1081	TATATTCCAGAGCTGGATGGACTCTCTGAGGAAAATAATGCTTGGTCCATCTCTACTAGT	1140
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1141	CGCTCCAGAGCTCGCCTGCAGATAGTGTGATCTCACCTGAATATTTATATCCAGTAGAA	1200
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1201	GATGATGTTATGAAAGGACAGACTCATTTGAGTCAGTCCCTGATTTCTGTGACGCCACTA	1260
Qy	421	IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe	440
Db	1261	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCATGACATCTTTTCATGTTTTT	1320
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Db	1321	CCCCAAGTCACGAAGAGGAAATTTGAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTC	1380
Qy	461	ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Db	1381	GTCATTTTATACAAAATTTACATCTATTTTAAAGAAAGCAATATAAACGATCCAGTGT	1440
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Db	1441	GGGCTGCTCTCCAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATATTACAGT	1500
Qy	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	1501	GGTGAATGGGAAGTTCTTGGCCGCAATGGATCTATATCAAGTTGATGAGTCAGAAG	1560
Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValSer	540
Db	1561	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCACT	1620
Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	1621	TACGTAAATCCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1680
Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	1681	ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACTGT	1740
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	1741	GTGTCCCTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTT	1800
Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	1801	TGGGCCACCAATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTTC	1860
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	1861	TCCTTTGAAGTACTACTCGGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA	1920
Qy	641	GlnProGlyLysGlyTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	1921	CAGCTCGAAAGAAATATCTTACTGTCTGTCTCATATATGGTGGTCTCTCAGTGCAGTTG	1980
Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	1981	GTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGTTGAATACCTTAGCTCTCTAGGT	2040
Qy	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	2041	TATGTGGTGTGATAGACAAACAGGGGATCTCTGTCCAGGGGCTTAATTTTGAAGGC	2100

Qy	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	Db	1698	ATCAGTCAGCACGTGACCTCTTTAAGCTAAGTAGTAGTAACACAGAAATCCACACTGT	1747
Db	608	GATCCAAAATATATGCCCTGCTGATCCAGACTCGGATTTGCTTTATACATAGCAACGATATT	667	Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLeuThrLysGluPhe	600
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	Db	1748	GTGTCCTTTTCAAGCTATCAGTCTCGAAGATGACCCCAACTTGCAAAACAAGGAAATTT	1807
Db	668	TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTA	727	Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260	Db	1808	TGGGCCACCACTTTGGATTTCAGCAGGTCTCTCTCTGACTACTACTCTCCAGAAATTTTC	1867
Db	728	GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAA	787	Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	Db	1868	TCCTTTGAAAGTACTACTCGGATTTACATTGATGGGATGCTCTACAGCCTCATGATCTA	1927
Db	788	TTTATAGATATCTCGCTATTTGGTGTGTCCTCCAAAGCTGAAACAACCTCCAGTGGTGT	847	Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	281	LysIleLeuArgIleLeuTyrGluAsnAspGluSerGluValGluIleIleHisVal	300	Db	1928	CAGCCTCGAAAGAAATATCTACTGTGCTTTCATATATGGTGGTCCCTCAGGTGCAGTGT	1987
Db	848	AAAATCTTAGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTATCATGTT	907	Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	1988	GTGAATAATCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGT	2047
Db	908	ACATCCCCTATGTTGGAAACAAGAGGGCGAGATTCAATCCGTTATCTCTAACAACAGGTACA	967	Qy	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	2048	TATGTGGTGTAGTAGACACAACAGGGGATCCTGTACCAGGGGCTTAAATTTGAAGGC	2107
Db	968	SCAAATCCTAAAGTCATTTTAAGATGTCAAGAAATATGATTGATGCTGAAGAAAGGATC	1027	Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Db	2108	GCTTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGACTCCAATAT	2167
Db	1028	ATAGATGTCAATAGATAAGCACTAAATTCACACCTTTTGAGATTTCTATTGAAGGAGTTGAA	1087	Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Qy	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	Db	2168	CTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGGGCATCCACGGCTGGTCTAT	2227
Db	1088	TATATTGCCAGAGCTGGATGACTCCTGAGGAGAAATATGCTTGGTCCATCTCTACTAGAT	1147	Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnSerSerAspIlePheArgValAlaIle	760
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Db	2228	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAAGATCTTCAGGGTTGCTATT	2287
Db	1148	CGCTCCAGAGCTCGCTGCAGATAGTGTGATCTCACTGAAATATTATTCACAGTAGAA	1207	Qy	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Db	2288	GCTGGGGCCCCAGTCACCTCTGGGATCTTCTATGATACAGATACACGGNACGTTATATG	2347
Db	1208	GATGATGTTATGGAAGGACAGAGACTCAATTGAGTCAGTGCTGATTCGTGAGCCACTA	1267	Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Qy	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	Db	2348	GGTCACCTCGACCAAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2407
Db	1268	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTT	1327	Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	Db	2408	AAAGTTCCTCTTGAAACCAAAATCGTTTACTGCTCTTACTATGTTTCTCGGATGAGAATGC	2467
Db	1328	CCCCAAAGTCACGAAGAGAAATTGAGTTTATTTTGCCTCGATTCGAACACAGGTTTC	1387	Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	Db	2468	CATTTTGCACATACCAGTATATATTCTAGTTTTTTTTAGTGAGGGCTGGAAAGCCATATGAT	2527
Db	1388	CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATTAACGATCCAGTGT	1447	Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500	Db	2528	TTACAGATCTATCTCTCAGAGAGACACACAGATAAGAGTTCTCGAATCGGAGAACATTAT	2587
Db	1448	GGGCTCGCTCGCTCCAAAGTATTCAAGTGTCTTATCAAGAGGAGATAGCAATACCACT	1507	Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
Qy	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	Db	2588	GAACTCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGATTTGCTGCTCTAAAA	2647
Db	1508	GGTGAATGGCAAGTCTTTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAG	1567	Qy	881	ValIle	882
Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540	Db	2648	GTGATA	2653
Db	1568	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1627	RESULT 3			
Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	AR631278			
Db	1628	TACGTAAATCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1687	LOCUS			
Qy	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	DEFINITION			
				Sequence 2 from patent US 6844180.			
				AR631278			
				ACCESSION			
				AR631278.1			
				VERSION			
				GI:59770920			
				DNA			
				linear			
				PAT 14-FEB-2005			
				2671 bp			

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 2671)

AUTHORS Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.

TITLE Serine protease genes related to DPPIV

JOURNAL Patent: US 6844180-A 2 18-JAN-2005;

Ferring BV,;

NLX;

FEATURES

source Location/Qualifiers

1..2671

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 2671
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-825-632-1 (1-882) x AR631278 (1-2671)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 Db 8 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTGGGACTGT 67
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 68 GAGGAGATATTGATATCACAGATCGGCTAAATGGAGCTTTTATGTGTGACGGTAT 127
 Qy 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 128 TCCTGGAGTCAGCTAAAGCTGCTTCCGATACCAAGAAATATCATGCTACATGATG 187
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 188 GCTAAGGCACCACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGCACCTCATTC 247
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 248 GACAGAATCTATTACCTTCCATGCTGTGAGAACAGAGAAATACACTGTTTATTTCT 307
 Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 308 GAAATTTCCAAACTATCAATAGACGACGCTTAAATGCTCTCTTGGAGCCCTCTTTTG 367
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 Db 368 GATCTTTTTCAGCAACTGACTATGGAATGATTTCTCGAGAGAGAACTATTTAAGA 427
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 428 GAAAGAAACGCAATGGACACATCGAATGCTCTTACGATTATCACCAAGGAAGTGA 487
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 488 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATCACGTAAAGATGAGGGGCCACAGA 547
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 548 TTATCCCAACCACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCACATACGATG 607
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 608 GATCCAAAATTTATGCCCTGCTGATCCAGCTGGATGCTTTATACATACACGATATT 667
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 Db 668 TGGATATCTAAACATCGTAACCAAGAGAAAGAGAGACTCTACTTATGTGCAATAGACTA 727

Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 728 GCCAATCATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAA 787
 Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 Db 788 TTTGATAGATATTCTGGCTATTGGTGTCTCCAAAGCTGAAACAACTCCAGTGGTGGT 847
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 Db 848 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATTCATGTT 907
 Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 908 ACATCCCTTATGTTGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 967
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 968 GCMAATCTTAAGTCACTTTTAAGATGTGAGAAATATATGATGATGCTGAAGGAGGATC 1027
 Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1028 ATAGATGTCATAGATAGGAACCTAATTCACCTTTTGAGATTCTATTTGAAGGAGTTGAA 1087
 Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1088 TATATTGCCAGAGCTGGATGGACTCTCGAGGAAATATATGCTGCTCCATCTACTAGAT 1147
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1148 CGCTCCAGACCTCGCTCGCAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAA 1207
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1208 GATGATGTTATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1267
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1268 ATTATCTATGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1327
 Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1328 CCCCAAGTCAAGAGAGAAATTCAGTTATTTTGGCTCTGATGCAAAACAGGTTTC 1387
 Qy 461 ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1388 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGT 1447
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysLysGluGluIleAlaIleThrSer 500
 Db 1448 GGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTATATCAAGAGGAGATAGCAATTTACCAGT 1507
 Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 Db 1508 GGTGAATGGGAAGTTCTTGGCCGCGATGATCTAATATCCAAAGTGTGATGAAGTCAAGAGG 1567
 Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
 Db 1568 CTGGTATATTTTGAAGGACCAAGAGCTCCCTTTTAGAGCATACCTGTAGTGTAGTCACT 1627
 Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 Db 1628 TAGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTCG 1687
 Qy 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 Db 1688 ATCAGTCAAGCACTGTGACTTCTTTATAGTAAGTATAGTACCAGAGAGAAATCCACACTGT 1747
 Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
 Db 1748 GTGTCCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGTGCAAAACAAAGAAATTT 1807

```
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1808 TGGGCCACCAATTTTGGATTTCAGAGGCTCTCTCTCGACTATACCTCCAGAAATTTTC 1867
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1868 TCTTTTCAAAAGTACTACTGGATTATACATTGTATGGGATGCTCTACAGACCTCATGACTA 1927
QY 641 GlnProGlyLysIleValProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1928 CAGCCTGGAAAGAAATATCTACTGTCTGTTCATATATGTGGTCTCAGTGCAGTTG 1987
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1988 GTGAATATCGATTTAAAGGAGTCAAGTATTTTCGGCTTGAAATACCTAGCTCTCTAGGT 2047
QY 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2048 TATGTGGTTGTAGTGATAGACAACAGGGGATCTGTCCACGAGGGCTTAATTTGAAGGC 2107
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2108 GCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGTGTGAAGGACTCCATAT 2167
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2168 CTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT 2227
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2228 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATT 2287
QY 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2288 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATG 2347
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2348 GGTCAACCTGTACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2407
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2408 AAGTTCCTCTCGAACCAATCGTTTACTCTTATACCTGTTTCTTGGATGAGAATGTC 2467
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2468 CATTTTGCATACATACAGTATATTACTGAGTTTTTTTAGTGGGGCTGGAAGGCCATATGAT 2527
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2528 TTACAGATCTATCTTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTAT 2587
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2588 GAACTGCACTCTTTTGGCACTACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTCTAAA 2647
QY 881 ValIle 882
Db 2648 GTGATA 2653

RESULT 4
AX608725 LOCUS 2671 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 2 from Patent WO0231134.
ACCESSION AX608725
VERSION AX608725.1 GI:28404298
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE 1
```

```
AUTHORS Qi,S., Akinsanya,K.O., Riviere,P.J. and Junien,J.L.
TITLE Novel serine protease genes related to appiv
PATENT: WO 0231134-A 2 18-APR-2002;
FERRING BV (NL)
FEATURES
    source
        1..2671
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2671
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-825-632-1 (1-882) x AX608725 (1-2671)
QY 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 8 ATGGCAGCAGCAATGGAACACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 67
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 68 GAGGAGATATTGAATCACAGGATCGGCCTAATTTGGAGCCITTTTATGTTGAGCGGTAT 127
QY 41 SerTyrSerGlnLeuLysLysLeuLeuAlaAspThrArgIleTyrHisGlyTyrMetMet 60
Db 128 TCCTGGATCAGCTTAAAGAGCTCTTGGCGATACCGAANAATATCATGGCTACATGATG 187
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 188 GCTAAGGACCAACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCAITCA 247
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 248 GACAGAACTTATACCTTGGCCATCTCTGGTGAGAACAGAGAAAATACACTGTTTATCTT 307
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 308 GAAATTCGCAAAACATCATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGAGCCCTTTTG 367
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 368 GATCTTTTTCAGGCAACACTGCACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAGA 427
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 428 GAAAGAAAACGCAATTTGGAAACAGTCGGAATTTCTTACGATTATCACCAGGAAGTGA 487
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 488 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGCTTAAAGATGAGAGGCCACAAGGA 547
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 548 TTTACGCAACAACCTTTAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCACATACCGATG 607
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 608 GATCCAAAATATATGCCCTGCTGATCCAGCTGGATGCTTTTATATACCAACAGATATT 667
QY 221 TrpIleSerAsnIleValThrArgGluGluArgIleuThrTyrValHisAsnGluLeu 240
Db 668 TGGATATCTTAACATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCAATGAGCTA 727
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 728 GCCACATGGAAGAAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTGTTCTCCAGAAGAA 787
```

QY	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Db	788	TTTGATAGATATTCTGGCTATTGGTGTCTCCAAAGCTGAAACAACTCCAGTGGTGGT	847
QY	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
Db	848	AAAATTCTTAGAATTCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTTCATGTT	907
QY	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Db	908	ACATCCCTATGTGGAAACAGAGGAGGAGATTCATTCGGTTATCTCTAAACAGGTACA	967
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Db	968	GCAATCTCTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATTGATGCTCAAGGAAGATC	1027
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Db	1028	ATAGATGTCATAGATAGGAACTAAATTCACACCTTTTGAGATTCTATTGGAAGGAGTTGAA	1087
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
Db	1088	TATATTGCCAGAGCTGGATGGACTCTCTGAGGNAATATGCTGGTCCATCTCTACTAGAT	1147
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1148	CGCTCCAGACTGGCTGCAGATAGTGTGATCTCACCTGGAATATTATTATCCAGTAGAA	1207
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1208	GATGATGTTATGAAGAGGAGAGACTCATTTGAGTGCCTGCTGATTCGTGAGCGCACTA	1267
QY	421	IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe	440
Db	1268	ATTATCTATGAAGAAACACAGACAICTGGATTAATATCATGACATCTTTCATGTTTTT	1327
QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Db	1328	CCCCAAGTCAAGAGGAGAAATTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1387
QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Db	1388	CGTCATTTATACAAATTTACATCTATTATTTAAAGAAAGCAAAATATAAACGATCCAGTGGT	1447
QY	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
Db	1448	GGGCTGCCTGCTCCAGTGATTTCCAGTGCTCTTCAAGAGAGATAGCAATTACCACT	1507
QY	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	1508	GGTGATGGGAAGTCTTTGGCCGGCATGGATCTAATATCAAGTTGATCAAGTCAAGAGG	1567
QY	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
Db	1568	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTGAAGCATCACCTGTACGTAGTCACT	1627
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	1628	TACGTTAAATCTCGAGAGGTGACAAAGCTGACCGCTGCTACCTACATCTTCTGCTGC	1687
QY	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	1688	ATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1747
QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	1748	GTGTCCTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAGAAACAAAGGAATTT	1807
QY	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	1808	TGGGCGACCATTTTGGATTCACGAGGTCCTCTTCCCTGACTATATCTCTCCAGAAATTTTC	1867
QY	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	1868	TCCTTTGAAAGTACTACTGGATTACATTGATGGGATGCTCTACAGGCTCATGATCTA	1927
QY	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	1928	CAGCCTCGAAAGAAATATCTCTACTGTCTGTTCATATATGCTGCTCCTCAGGTGCAGTTG	1987
QY	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	1988	GTGAATAATCAATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGGCTCTCTAGGT	2047
QY	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	2048	TATGTGGTTGTAGTGATAGACACAGGGGATCTGTCTCCAGGGCTTAAATTTGAAGGC	2107
QY	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	2108	GCCTTTAAATATAAATGGGTCAAATAGAAATTCAGCATCAGGTGAAGGACTCCAAATAT	2167
QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Db	2168	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATCCACGGCTGGTCTAT	2227
QY	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2228	GGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATT	2287
QY	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2288	GCTGGGGCCCCAGTCACCTCTCTGGATCTTCTATGATACAGATACACGGAACGTTATATG	2347
QY	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	2348	GGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2407
QY	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2408	AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTC	2467
QY	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2468	CATTTTGACATACACAGTATATTACTAGATTTTTTAGTGGGCTGGAAAGCCATATGAT	2527
QY	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2528	TTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTAT	2587
QY	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
Db	2588	GAACTGCATCTTTTGCACTTACCTTCAGAAACCTTGGATCACGTATTTGCTCTTAAAA	2647
QY	881	ValIle 882	
Db	2648	GTGATA 2653	
RESULT 5			
AY354202			
LOCUS			
DEFINITION			
Homo sapiens dipeptidyl peptidase 8 isoform 3 (DPP8) mRNA, complete cds, alternatively spliced.			
ACCESSION			
AY354202			
VERSION			
AY354202.1 GI:34329351			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Cattaehini;			
Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 3030)			
AUTHORS			
Sha, J.H., Zhou, Z.M. and Li, J.M.			
TITLE			
Direct Submission			
JOURNAL			
Submitted (29-JUL-2003) Key Lab of Reproductive Medicine, Nanjing Medical University, 140 Han Zhong Road, Nanjing, Jiangsu 210029,			

China	Location/Qualifiers																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
-------	---------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 1709 GGTGAATGGGAAGTTCTTGGCGCATGGATCTAATATCCAAGTTGATCAAGTCAAGAAG 1768
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1769 CTGGTATATTTTGAAGGACCAAGAGCTCCCTTTTGAAGCATCACCTGTACGTAGTCAGT 1828
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1829 TACGTAAATCTCGAGAGGTGACAGGCTGACTGACGCTGGCTACTACATCTTCTGCTGC 1888
Qy 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1899 ATCAGTCAGCACTGTGACTCTTTATTAAGTAAAGTATAGTAAACCAAGAAATCCACACTGT 1948
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1949 GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGATCCCAACTTGCACAAACAAAGGAATTT 2008
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2009 TGGGCGACCAATTTTGGATTACAGCAGGTCTCTCTGACTATATCTCTCCAGAAATTTTC 2068
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2069 TCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTACAAAGCTCATGATCTA 2128
Qy 641 GlnProGlyLysLysTyrProThrValIlePheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2129 CAGCCTGGAAAGAAATATCTTACTGTCTGTCTCATATATGTTGGTCTCTCAGGTGCAGTTG 2188
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2189 GTGAATAATCGGTTTAAAGAGTCAAGTATTTTCGGTTGAATACCCCTAGCCTCTCTAGGT 2248
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2249 TATGTGGTTGTAGTATAGACACAGAGGATCTGTCTCACCAGGGCTTAAATTTGAAGGC 2308
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyr 720
Db 2309 GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGAAGGACTCCAAATAT 2368
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2369 CTAGCTCTCGATATGATTTTCATTTGACTTGTGATGATGATGATGATGATGATGATGAT 2428
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2429 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTCAGGGTTCGTATTT 2488
Qy 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2489 GCTGGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGATACAGGAAAGCTTATATG 2548
Qy 781 GlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2549 GGTCAACCTTGACCAAGTAAACAGGGCTATTCTTAGGATCTGTGGCCATGCAGCAGAA 2608
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2609 AAGTTCCCTCTGAACCAAAATCGTTTACTGTCTTACATGGTTTCTGGATGAGATGTC 2668
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2669 CATTTTGCATACACAGTATATTTACTGAGTTTTTTTGTAGGGCTGGAAGCCATATGAT 2728
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2729 TTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGGAGAACATTAT 2788
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880

Db 2789 GAACGTGATCTTTTGCACACTACCTTCAGAAAAACCTTGCATCAGTATTGTGCTTAAAA 2848
Qy 881 ValIle 882
Db 2849 GTGATA 2854
RESULT 6
AX342633 3106 bp DNA linear PAT 12-JAN-2002
LOCUS Sequence 30 from Patent WO0198468.
DEFINITION AX342633
ACCESSION AX342633
VERSION AX342633.1 GI:18152030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Elliott,V.S., Gandhi,A.R., Lal,P., Au-Young,J.,
Tribouley,C.M., Deleageane,A.M., Baugn,M.R., Nguyen,D.B., Lee,E.A.,
Hafalia,A., Khan,F.A., Walia,N.K., Yao,M.G., Lu,D.A., Patterson,C.,
Tang,Y.T., Walsh,R.T., Azimzai,Y., Ramkumar,J., Xu,Y. and Reddy,R.
JOURNAL Patent: WO 0198468-A 30 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source 1..3106
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="incyte ID No: 7160544CBI"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3106
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-825-632-1 (1-882) x AX342633 (1-3106)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 203 ATGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTGAAACCTGGCGACTGT 262
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 263 GAGGAATATTGAATTCACAGGATCGGCTTAATTTGGAGCCCTTTTATGTGACGGTAT 322
Qy 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 323 TCTGTGAGTCAGCTTAAAGAGCTGTTCGGATACCCAGAAATATCATGGCTACATGATG 382
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 383 GCTAAGGCCACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCAATCA 442
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 443 GACAGAATCTATTACCTGCCATGTCGTGTGAGAACAGAGAAATATACACTGTTTATTCT 502
Qy 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTyrLysProLeuLeu 120
Db 503 GAAATTCCTCCAAATCATCAATAGCAGCAGCTCTTAATGCTCTCTTGGAGCCCTTTTGT 562
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140
Db 563 GATCTTTTTCAGCAACACTGGACTATGGAATGTTATCTTCGAGAGAGAACTATTAGA 622
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

Db	623	GAAGAAAAACGATTTGGAACAGTCGGAATTCCTTACGATTATTCACCAAGGAAGTGGGA	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	1763	CTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTCACT	1822
Db	683	ACATTTCTGTTTCAAGCCGAGTGTGAATTTATACGTAAAGATGGAGGGCCCAAGGA	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	1823	TACGTAAATCTCTGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1882
Db	743	TTTACGCAACAACCTTTTAAGGCCCAATCTAGTGGAAACCTAGTTGTCCCAACATACGGATG	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	201	AspProLysLeuCysProAlaAspProAspTrrPileAlaPheIleHisSerAsnAspIle	1883	ATCAGTCAGCACTGTGACTTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCCACTGT	1942
Db	803	GATCCAAAATATATGCCCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	1943	GTGTCCTCTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAGAAACAAAGAAATTT	2002
Db	863	TGGATATCTTAACATCGTAACAGAGAAAGAGAGACTCACTATATGTGCACAATGAGCTA	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	2003	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTCGACTATATCTCTCCAGAAATTTTC	2062
Db	923	GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	261	PheAspArgTyrSerGlyTyrTrrTrrCysProLysAlaGluThrThrProSerGlyGly	2063	TCCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGGCTCATGATCTA	2122
Db	983	TTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAACTCCCGAGTGGT	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	2123	CAGCTCGAAAGAAATATCTTACTGTCTGTTCATATATGCTGTCTCTCAGGTGCAAGTTG	2182
Db	1043	AAATTTCTTAGAATCTTATTAAGAAATAATGATGAATCTGAGGTGGAATAATATTCATGTT	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	2183	GTGAATATTCGGTTTAAAGGATCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGT	2242
Db	1103	ACATCCCCCTATGTTGGAACCAAGAGGGCAGATTCATTCGTTATCTCTAAACAGGTACA	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	2243	TATGTGGTTGTAGTGATAGACACAGGGGATCTGTGCCAGCGGCTTAAATTTGAAGGC	2302
Db	1163	GCAATCTCTAAAGTCACCTTTTAAGATGTCCAGAAATAATGATGATGCTGAGGAAGGATC	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	2303	GCCTTTTAAATATAAAATGGGTCAATAGAAATAGACATCAGGTGGAGGACTCCAATAT	2362
Db	1223	ATAGATGTATAGATAAGAACTAAATCAACCTTTTGGAGATCTATTTGAAGAGAGTTGAA	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrrSerTyr	740
Qy	361	TyrIleAlaArgAlaGlyTrrThrProGluGlyLysTyrAlaTrrSerIleLeuLeuAsp	2363	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGCTGTCTAT	2422
Db	1283	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAAATATGCTTGGTCCATCTCTAGAT	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	2423	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAAGTATCTTTCAGGGTTGCTATT	2482
Db	1343	CGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCTGAAATTTATTCCTCAGTAGAA	761	AlaGlyAlaProValThrLeuTrrPilePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	2483	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2542
Db	1403	GATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTGCCTGATCTGTGACGCCACTA	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Qy	421	IleIleTyrGluGluThrThrAspIleTrrPileAsnIleHisAspIlePheHisValPhe	2543	GGTCACCTTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGGCATGCAAGCAGAA	2602
Db	1463	ATTATCTATGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTCATGTTTTT	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	2603	AGTTTCCCTCTGAACCAATTCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAAATGTC	2662
Db	1523	CCCCAAAGTCAAGAGGAAGAAATTTAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTrrLysArgSerGly	2663	CATTTTGCACATACCAGTATATTTACTGATTTTTTTAGTGGGCTGGAAGCCCATGAT	2722
Db	1583	CGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	2723	TTACAGATCTATCTCCAGAGACACAGCATAGAGTTCTGTAATCGGAGAACATTAT	2782
Db	1643	GGGCTGCCCTGCCAAGTGAATTTCAAGTGTCTTCAAGAGAGAGATAGCAATATCCAGT	861	GluLeuHisLeuLeuHisTyrIleGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Qy	501	GlyGluTrrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	2783	GAACTGCATCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCAGCTATTTGCTGCTCTANAA	2842
Db	1703	GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTTAATATCCAAGTTGATGAATCAGAAAG			


```
QY      881 ValIle 882      |||||
Db      2843 GTGATA 2848      |||||

RESULT 7
AR651452
LOCUS      AR651452      3120 bp      DNA      linear      PAT 20-APR-2005
DEFINITION Sequence 2 from patent US 6881564.
ACCESSION AR651452
VERSION    AR651452.1 GI:62795938
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 3120)
AUTHORS   Abbott, C.A. and Gorrell, M.D.
TITLE     Dipeptidyl peptidases
JOURNAL   Patent: US 6881564-A 2 19-APR-2005;
          The University of Sydney; Sydney;
          AUX;

FEATURES   Location/Qualifiers
            source      1..3120
                        /organism="unknown"
                        /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3120
Score:          4700.00      Matches:      882
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      100.0%      Indels:      0
DB:              6      Gaps:          0

US-10-825-632-1 (1-882) x AR651452 (1-3120)
QY      1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db      214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAATCTGGGACTGT 273
QY      21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db      274 GAGGAGAAATATGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTGAGCGGTAT 333
QY      41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db      334 TCCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAGAAATATCATGCTACATGATG 393
QY      61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db      394 GCTAAGGCACCAATGATTTTCATGTTGTGTAAGAGGAATGATCCAGATGGACCTCATTTCA 453
QY      81 AspArgIleTyrTrpLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db      454 GACAGAAATCTATTACCTTCCCATGTCTGTGAGAACAGAGAAATACACTGTTTATTTCT 513
QY      101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db      514 GAAATTTCCCAAAATCATCAATAGACAGCAGCTTTAATGCTCTCTTGGAAAGCCCTTTTGG 573
QY      121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db      574 GATCTTTTTCAGGCAACACTGGCACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGA 633
QY      141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db      634 GAAGAAAGACCATTCGAACAGTCGGAATGCTTCTTACGATTATCACCAAGGAAGTGA 693
QY      161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db      694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAGATGGAGGCGCCCAAGGA 753
QY      181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

|||||
754 TTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCACATACCGGATG 813
QY      201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db      814 GATCCAAAATATATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
QY      221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
Db      874 TGGATATCTTAACATCGTAACAGAGAGAGAGAGACTCACTTATGTGCACATAGACTA 933
QY      241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db      934 GCCAACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAGAA 993
QY      261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db      994 TTTGATAGATATTTCTGGCTATTTGGTGTCTCAAAAGCTGAACAACACTCCCGTGGTGT 1053
QY      281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db      1054 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTTCATGTT 1113
QY      301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db      1114 ACATCCCCCTATGTTGGAACAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
QY      321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db      1174 GCAATCTCTAAGTCACTTTTAAGATGTCAAGAAATATGATGATGCTGAAGGAGGATC 1233
QY      341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db      1234 ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAGATTCCTATTGAAGGAGTTGAA 1293
QY      361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db      1294 TATATTGCCAGAGCTGATGGACTCTCTGAGGAAATATGCTTGGTCCATCTACTAGAT 1353
QY      381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db      1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTATCCCACTAGTAA 1413
QY      401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db      1414 GATGATGTTATGGAAGAGCAGAGACTCATTTGAGTCAGTGCCTGATCTGTGACCCACTA 1473
QY      421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db      1474 ATTATCTATGAGAAACAACACAGACATCTGGATTAATATCCATGACATCTTTCATGTTT 1533
QY      441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db      1534 CCCCNAAGTCACGAGAGGAATTCAGTTATTTTTCCTCTGAATGCAAAACAGGTTTC 1593
QY      461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db      1594 CGTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAAAATATAAACCATCCAGTGGT 1653
QY      481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db      1654 GGGCTGCTGCTCTCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
QY      501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db      1714 GGTGAATGGGAAGTCTTGGCCGCGATGGATCTAATATCCAAAGTGTGATGAAGTCAGAAGG 1773
QY      521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db      1774 CTGTATATTTTGAAGGCAACCAAGACTCCCTTTTAGAGCATCACCTGTAGTACGTAGTCA 1833
QY      541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
|||||
```

Db	1834	TACGTAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGCTACTCACATTTCTTGCTGC	1893
Qy	561	ILeSerGlnHisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCys	580
Db	1894	ATCAGTCAGCACTGTGACTTCCTTTATAGTAAGTAGTAGTAACCAAGAAATCCACACTGT	1953
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	1954	GTGTCCCTTTACAAGCTATCATAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGGAATTT	2013
Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	2014	TGGCCACCACTTTGATTCAGCAGGTCTCTCTCTGACTACTCTCTCCAGAAATTTTC	2073
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	2074	TCCTTTGAAAGTACTACTGGAATTTACATTTATGGAATGCTCTACAGCCTCATGATCTA	2133
Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu	660
Db	2134	CAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCAGTGCAATG	2193
Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	2194	GTGAATAATCGTTTAAAGAGTCAAGATATTTCCGCTTGAATACCTAGCCTCTCTAGGT	2253
Qy	681	TyrValValValValIleAspAsnA:gglySerCysHisArgGlyLeuLysPheGluGly	700
Db	2254	TATGTGGTTGTAGTGATAGACAACAGGGATCTCTGTACCGAGGCTTTAAATTTGAAAGC	2313
Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	2314	GCCTTTAAATATAAATGGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCAATAT	2373
Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Db	2374	CTAGCTTCTCATATGATTTTCATTTAGCTTAGATCGTGTGGCATCCAGGCTGGTCTAT	2433
Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2434	GGAGATACCTCTCCGTATGGCATTAATGCAGAGTCAATATCTTCAGGGTTGCTATT	2493
Qy	761	AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrClnArgTyrMet	780
Db	2494	GCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGTTATAG	2553
Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetClnAlaGlu	800
Db	2554	GGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAACAGAA	2613
Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2614	AAAGTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATG	2673
Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2674	CATTTTGCACATACACAGTATATTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGAT	2733
Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2734	TTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCTCGAATCGGAGAAACATTAT	2793
Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaLeuLys	880
Db	2794	GAAGTGCATCTTTTGCACTACCTTCAGANAAACCTTGGATCACGTTATGCTGCTCAAAA	2853
Qy	881	ValIle 882	
Db	2854	GTGATA 2859	
RESULT 8			
AF221634			
LOCUS			
Db	AF221634	3127 bp	mRNA linear PRI 05-NOV-2000
Qy	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	214	ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGAACTCGGACTGT	273
DEFINITION	Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, complete cds.		
ACCESSION	AF221634		
VERSION	AF221634.1		
KEYWORDS	GI:11095187		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 3127)		
AUTHORS	Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W. and Gorrell,M.D.		
TITLE	Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8		
JOURNAL	Eur. J. Biochem. 267 (20), 6140-6150 (2000)		
PUBMED	11012666		
REFERENCE	2 (bases 1 to 3127)		
AUTHORS	Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver Centre, Centenary Institute of Cell Biology and Cancer Medicine, Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia		
FEATURES	Location/Qualifiers		
source	1..3127		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="15"		
	/map="15q22"		
gene	1..3127		
	/gene="DPP8"		
CDS	214..2862		
	/gene="DPP8"		
	note="serine protease; member of prollyl oligopeptidase S9b family belonging to enzyme clan SC; alternatively spliced forms of this mRNA are found in AF221635, AF221636 and AF221637"		
	/codon_start=1		
	/product="dipeptidyl peptidase 8"		
	/protein_id="AAG29766.1"		
	/db_xref="GI:11095188"		
	/translation="MAAMETEQLGVEIFETADCEINIESQDRPKLEDFYVERYSWSQ LKLLADTRKYHYGMMKAPHPDMFKRNDPQPHSDRIYILAMSGENRNTLFYSRI PKTINRAVLNLSKPLDLFOATLDYMYSRREELLRRKRKIGTVGLASDYTHQSG TFLFQAGSGYHVKGQPGQFTQOPLRNLVETSCNIRMDPKLPADPDWATFATHSN DIWISNIVTREERLTYVHNELNMEEDARSAGVATFVLOESBFDYRGVYKWCCKAEIT PSQKILRILYEENDESEVEI IHVTSPLMTRADSFYKGTGTANPKVTPKMSIEMI DAEGRIIDVDKELIQFEILFEGVEYIARAGWTEGKYAWSILLDRSQRLOQVLIS PELFIPVEDDDVMERQLIESVPDVTPLIYEETDINIHIHDFHVPQSHEEBIEF IPASECTGFRLYKITSILSKSKYKSSGGLPAPSDFKPIKEEITATSGSCEVILGR HGSINQKNPHCVLYSLKSSPEDDPTCKTEFWATILDSAGPLPDYTPPPIFSPES FFIISKYNGKPHDLQPKKYPTVLFYGGQVQLVNNRFGKVKYFRNLTLASLGYS VVVDNRSCHRGKLFEGAFKYMGOIEIDDOVEQLQYLASRYDFILDRVGIHWSY CGYLSLMALMORSDFRVALAGAPVTLMIFYDTGTYERYMGHPDQNEQGYLGSVAMQ AKKPSSENRLLLAHGLDENVHPAHTSILLSFLVRAGKPYDLQIYQPSRHSIRVPES GEHYELHLLHYLQENLGSRIAALKVI"		
ORIGIN			
Alignment Scores:	0	Length:	3127
Pred. No.:	4700.00	Matches:	882
Score:	100.0%	Conservative:	0
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	100.0%	Indels:	0
Query Match:	8	Gaps:	0
DB:			
US-10-825-632-1 (1-882) x AF221634 (1-3127)			

QY 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 274 GAGGAGATATATGATACAGATCGGCCTAAATTTGGAGCTTTTATGTTGAGCGGTAT 333
QY 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGACACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGAGCTCATTTCA 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAACTTATACCTTGGCCATGCTCTGGTGAGAACAGAAAATACATGTTTATTTCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTTCCAAACATATCATAGAGCAGAGCTCTTAATGCTCTCTTGGAGGCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
DB 574 GATCTTTTTCAGGCACACTGGACTATGGAAATGATTTCTCGAGAAGAAGACTATTAGA 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAACGCAATGGAAACAGTCCGAATTCCTTACGATTTATCACCAGGAAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAGAGTGGAGGCCACAGA 753
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTTACGCCAACACCTTTAAGGCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 814 GATCCAAATTTATGCCCCGCTGATCCAGACTGGATTCCTTTATACATAGCAACGATAT 873
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
DB 874 TGGATATCTTAACATCGTAACAGAGAGAGAGAGACTCACTTATGTGCACATGAGCTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 934 GCCAATCATGGAAGAGATGCCAGATCAGCTGGAGTCCCTACCTTTGTTCTCCAGAAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAGCTGAACACTCCAGTGGTGT 1053
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
DB 1054 AAAATTTCTTAGAATCTTATAGAGAAAATCATCAATCTGAGGTGGAAATATTATCATGTT 1113
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTTATGTTGAAACAGAGGAGGCAGATTCATTTCCGTTATCTCAAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 1174 GCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAAATATGATGATGCTGAAGAGGATC 1233
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCATAGATAAGAACTAATCAACTTTTGGATTCATTTTGAAGAGGTGAA 1293
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTTGCCAGAGCTGGATGGACTCTCGAGGGAATAATGCTTGGTCCATCTTACTAGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400

DB 1354 CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACCTGAAATATTATTTATCCAGTAGAA 1413
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
DB 1414 GATGATGTTATGGAAGGACAGAGACTCATTTGAGTCACTGCTGATGTTCTGTGACGCCACTA 1473
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1474 ATATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1534 CCCCAAGTCAAGAGAGGAAATTTGAGTTTATTTTGGCTCTGTAATCAAAACAGGTTTC 1593
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
DB 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAACATCCAGTGGT 1653
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
DB 1654 GGCTCGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTTACAGT 1713
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
DB 1714 GGTGAATGGAAATTTCTTGGCGCGCATGGATCTAATATCAAGTTGATGAAGTCAGAAG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGGCATCACTGTAGTAGTCACT 1833
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TAGTAAATCTCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTACCAAGAGAAATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCCCTTTACAAGCTATCAAGTCTCAAGATGACCCACTTCGCAAAACAAAGGAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
DB 2014 TGGGCCACCAATTTGGATTTGACAGGCTCTCTCTGACTATATCTCTCAGCAAAATTTTC 2073
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 2134 CAGCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATGTTGGTCTCTCAGTGCAGTTG 2193
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCTTAGCCTCTCTAGT 2253
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
DB 2254 TATGTGTTGTAGTAGACACCAAGGGGATCTGTCTCCAGGGGCTTAAATTTGAAGGC 2313
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
DB 2314 GCCTTTAAATATATAATGGGTCAAAATAGAAATTTGACATAGGTGGAGGAGCTCCAATAT 2373
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
DB 2374 CTAGCTTCTCGATATGATTTTCAATTTGACTTAGATGCTGGGGATCCAGCGCTGCTCTAT 2433
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760

Db	2434	GGAGGATACCTCTCCCTGATGGCAATTAATGACAGAGTCAGATATCTTCAGGGTGCCTATT	2493
Qy	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2494	GCTGGGCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2553
Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	2554	GCTCACCTGACCAAGATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAACGAGAA	2613
Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2614	AAGTCCCTCGTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC	2673
Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2674	CATTTTGCACATACACAGTATATTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGAT	2733
Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2734	TTACAGATCTATCTCTAGGAGACACAGCATAAAGATTCTCTGAAATCGGAGAACATTAT	2793
Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	2794	GAACCTGCACTCTTTGCACTACCTTCAAGAAAACTTTGGATCAGTATTGCTCTTAAA	2853
Qy	881	Vallie 882	
Db	2854	GTGATA 2859	
RESULT 9			
AX354793		3143 bp	DNA linear PAT 06-FEB-2002
LOCUS		Sequence 1 from Patent WO0179473.	
DEFINITION		AX354793	
ACCESSION		AX354793.1	GI:18619526
VERSION			
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1	Meyers, R. A. and Williamson, M.
AUTHORS			21953, a human prollyl oligopeptidase family member and uses thereof
TITLE			Patent: WO 0179473-A 1 25-OCT-2001.
JOURNAL			Millennium Pharmaceuticals, Inc. (US)
FEATURES		Location/Qualifiers	
source		1..3143	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
		229..2877	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD22990.1"	
		/db_xref="GI:18619527"	
CDS		/translation="MAAAMETEQLGVEIFETACBENIESQDRPKLEFPYVERYSWSQ LKLLADTRKYHYGMYMAKPHDFMVKRNDPDGPHSDRIYILAMGENRENTLFYSEI PKTINRAVLMSLWKLDFQATLDYGMYSREELLRERKRIGTVGASIDYHGSG TFLFOAGSGYVHVKGPGQFTQOPLRNLVETSCFNIRMDPKLCPADPDWTFIHSN DIWISNIVREERLTYVHNELANMEEDARSAGVATFVLOEEFDYRGYVWCPKAEIT PSGGKILRIYEENDESEVEIHTVSMLETRRADSFYKPTGTANPKVTFKMSIMI DAEGRIDVIDKEILQPEILFEGVEYIARAGWTPEGKAWSLIDRSQRLQIVLIS PELFTIPVEDDWMERQRLTISVPDSVPLIYEETDWINIHDIHDFVFPQSEEEIEF IPASCKTGFRRHLVKITSLKESKYKRSGLPAPSDFKCPDKIEAITSGEWEVLGR HFSNQVDEFRLLVYFEGTKDSPLEHHLVYVSVNPGEVTRLTDGYSHSCISQCHD PFISKYSNPKPHCVSLYKSPEDDPCTKFEFMATILDSAGPLPYDTPPEIFSFES TGTFTLYGMLYKPHLOPKKYPTVLFYGGPOVLNNRFKGVYFRNLTLASIGYV VVVIDNRGSHRGKLFKAGPKYKMGQIEIDQVEGLQYLASRYDFDIDRVDIGHWSY GVGLSLMALMORSDI FRVATAGPVTLWIFDYGTGYRYMGHPDQEQYVILGYSVAMQ AEKFPSEPNRLLLHGLFDENVHFAHTSILLSFLVRAGKPYDLQIYPQERHSRVPES GEHYELHLLHYLQENLGRINAULKVI"	

Db 1129 ACATCCCTATGTTGGAAACAGGAGGAGATTCATTCCTGTTATCCCTAAACAGGTACA 1188
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyValGile 340
Db 1189 GCAATCCTTAAGTCACTTTTAAGATGTGAGAAATATGATTTGATGCTGGAAGGAGATC 1248
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1249 ATAGATGTCATAGATAGGAACATAATCAACCTTTTGAGATCTTATTTGAAGAGGTTGAA 1308
Qy 361 TyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1309 TATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTCTAGAT 1368
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1369 CGCTCCAGACTCCCTCGAGATAGTGTGATCTCACCTGATTAATTTATCCAGTAGAA 1428
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1429 GATGATGTTATGGAAGAGCAGAGACTCATTGAGTCAGTGCCTGATTCGTGAGGCCACTA 1488
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1489 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1548
Qy 441 ProGlnSerHisGluGluGluIleGluPheIleAspLeuAspGluCysLysThrGlyPhe 460
Db 1549 CCCCAAGTCAAGAGAGGAAATGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC 1608
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1609 CGTCATTTATACAAATTTACATCTATTTTAAAGAAAGCAATATAAAGCATCCAGTGGT 1668
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 1669 GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCCTATCAAGAGAGAGATAGCAATACCAGT 1728
Qy 501 GlyGluTrpGluValLeuLysArgHisGlySerAsnIleGlnValAspGluValArg 520
Db 1729 GGTGAATGGGAAGTCTTGGCCGCATGGATCTAAATATCAAGTTGATGAAGTCAGAAG 1788
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
Db 1789 CTGGTATATTTTGAAGGACCAAGATCTCCCTTTAGAGCATCACCTGACGTAGTCAGT 1848
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1849 TAGCTAAATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1908
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1909 ATCAGTCAGCAGCTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1968
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1969 GTGTCCCTTTACAGCTATCAAGTCCCTGAGATGACCACTTGCAAAACAAAGNAATTT 2028
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2029 TGGGCCACCAATTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2088
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2089 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTA 2148
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyLysProGlnValGlnLeu 660
Db 2149 CAGCTCGGAAGAAATATCTTACTGTGCTGTCTATATGTTGTTCTCAGGTGCAAGTTG 2208
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680

Db 2209 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCGGCTTGAATACCTAGCCTCTCTAGGT 2268
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2269 TATGTGGTTGTAGTCATAGACAACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGC 2328
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2329 GCCTTTAAATATAAATGGTCAATAGAAATTCAGCATCAGGTGGAGGACTCCATAT 2388
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2389 CTAGCTTCTCGATATGATTTTCATTTAGATCGTGTGGCATCCACGGCTGTCTCTAT 2448
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2449 CGAGGATACCTCTCCCTGATGGCATTATGACAGAGGTTCAGATATCTTCAGGGTGTCTATT 2508
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2509 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGACGTTATATG 2568
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2569 GGTCACTCCCTGACAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAA 2628
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2629 AAGTTCCCTCTGAACCAATCGTTACTGTCTTACATGCTTTCTCTGGATGAGAATGC 2688
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2689 CATTTTGCACATPACAGATATATTTACTGTGATTTTGTAGGGCTGTGAAAGCCATATGAT 2748
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2749 TTACAGATCTATCTCCAGAGAGACACAGATAGAGTCTCTGAATCGGAGAACATTA 2808
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2809 GAATGTCATCTTTTGCACTACCTTCAGAAAACCTTGGATCAGCTATTGCTGCTCTAAAA 2868
Qy 881 ValIle 882
Db 2869 GTGATA 2874

RESULT 10

BC040203
LOCUS
DEFINITION
IMAGE:4824813, complete cds.
ACCESSION
BC040203
VERSION
BC040203.1
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 4535)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, I.B., Ioshiki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abrahamson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

BC040203 4535 bp mRNA linear PRI 07-OCT-2003

Homo sapiens dipeptidylpeptidase 8, mRNA (cdna clone MGC:42570

IMAGE:4824813), complete cds.

ACCESSION BC040203

VERSION BC040203.1

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1 (bases 1 to 4535)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Scapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, I.B., Ioshiki, S.,

Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,

Abrahamson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Bartels, Y.S., Kozminski, M.I., Skalek, U., Smaluk, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2. (bases 1 to 4535)
Strausberg, R.
Direct Submission
Submitted (27-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunsatane, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nananavi,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 71 Row: k Column: 23.
Location/Qualifiers
1. .4535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:42570 IMAGE:4824813"
/tissue_type="Testis"
/clone_lib="NIH MGC 97"
/lab_host="DH10B"
/note="Vector: pBluescript"
1. .4535
/gene="DPP8"
/note="synonyms: FLJ20283, FLJ14920, MGC26191"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"
1629..4277
/gene="DPP8"
/codon_start=1
/product="DPP8 protein"
/protein_id="AAH40203.1"
/db_xref="GI:26007917"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"
/translations="MAAAMETSQLGVEIFETADCEENIESQDRPKLPFFVVERYSWQ
LKLLADTRYHYGMKAPDFMFKVNRDPGSHSDRIIYLAHSGNSRNTILFYSEI
PKTINRAADLMSKPLLDLFOALDYGMYREELRKRIGTVGLASVDYHQSGD
TFIFQAGSGYHYKDGQPGQTQQPLRPNLVETSCFNIMDPKPCPADPWPFIHGN
DIWISNIVREERLTYVHNELANMEEDARSAGVAIFVLQEEFDYSGYMWCPKAET
PSGKILRLIYEENDESEVEIHHVTSPLMLETNRADSPRYKGTANPKVTFKMEIMI
DAGRIIDVLDKLIQPFILPEGYETARAGWTEPKYAWISILLDRSQRLOLVIS
PELPIVEDDWMQRQLIESVDSVTPLIYYEETDDIWINTHIDIFHVPFQSHHEEILF
IFASECTDPRHLRYKITSILKESKYKSGGLPAPSPDKPCIKBEIAITSGWEVLGR
HGNIQVDEVRNVYFEGTKDSPLEHLVYVSNVNPGEVRLTDGRYSHSCISQCHD
FFTSKYSGNPKHCVSLYKLSGSPDEDDTKTFEWTALDSAGPLPYTPPEIFSPFS
TTGFTLYGMLYKPHDLQPGKYPVLFVIYGGPQVQVNNRFPKGVYFRLNLTASLGVV
VVIDNRGSHRGLKFEQAFKYMKGQIEIDQVEGLVLAISRYDFILDLRVGIHGSY
GGVLSLMALMQRSDIFRVAIAGAPVTLWTFYDTGYTERYMGHPQNEGYVLSVAMQ
AEKFPSPENRLLLLHGFLDENVHFAHTSILLSLFLVRAGKPYDLQIYPOERHSRVPES

misc_feature		GEHYELHLHYLQENLGSRIAALKVI" 3207..4196 /gene="DPP8" /note="DAP2: Region: Dipeptidyl aminopeptidases/acylaminoacyl-peptidases [Amino acid transport and metabolism]" /db_xref="CDD:COG1506"	
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	4535
Score:	4696.00	Matches:	881
Percent Similarity:	99.9%	Conservative:	0
Best Local Similarity:	99.9%	Mismatches:	1
Query Match:	8	Indels:	0
DB:	8	Gaps:	0
US-10-825-632-1 (1-882) x BC040203 (1-4535)			
QY	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
DB	1629	ATGGCAGCAGCAATGGAAACAGACACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT	1688
QY	21	GluGluAenIleGluSerGlnAspArgProLyLeuGluProPheTyrValGluArgTyr	40
DB	1689	GAGGAGAATAATTGAATCACAGGATCGGCCTAAATTGGAGCCTTTTATGTGTGAGCGGTAT	1748
QY	41	SerTTPSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
DB	1749	TCTTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACCAGAAATAATCATCGCTACATGATG	1808
QY	61	AlaIleValProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
DB	1809	GCTAAGGACCAACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGTGCACCTCATTTCA	1868
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
DB	1869	GACGAATCTAATACCTTGGCATGCTGGTGAGACAGAGAAATACTACTGTTTATTTCT	1928
QY	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
DB	1929	GAATTTCCAAAACATATCAATAGACGACGAGTCTTAATGCTCTCTTTGGAAGGCTCTTTTG	1988
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
DB	1989	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAAGAAGAAGAACTATTAAAG	2048
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
DB	2049	GAAGAAAAACGCAATGGACAGTCGGAATTTGCTTTTACGATTATCACCAGGAAGTGGG	2108
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
DB	2109	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCCACAAAGGA	2168
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
DB	2169	TTTACGCAACCACTTTTAAGCCCACTCTAGTGGAAACTAGTTGTCCTCCCACTACGGATG	2228
QY	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
DB	2299	GATCCAAAATATCCCTGCTGATCCAGCTGGATGCTTTTATATCATCGCAACGATATT	2288
QY	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnLeuLeu	240
DB	2289	TGGATATCTAAATCGTAACACAGAGAAAGAGAGACTCACTTATGTGCACAAATAGCTA	2348
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
DB	2349	GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTGCCTTCTTCTTCCAAAGAA	2408
QY	261	PheAspArgTyrSerGlyTyrTrpTTPCysProLysAlaGluThrThrProSerGlyGly	280

Db 2409 TTTGATAGATATTCCTGGCTATTTGGTGTGTCCTCAAAAGCTCAAAACATCTCCAGTGTGTGT 2468
QY 281 LysileLeuArgileLeuTyrGluGluAsnAspGluSerGluValGluIleleHisVal 300
Db 2469 AAAATCTTAGAATTTCTATATGAAGAAATGATGATCTGAGTGGAAATTTATCTATGTT 2528
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 2529 ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCGGTTATCTTAAACAGGTACA 2588
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 2589 GCAATCTCTAAAGTCACATTTTAAAGATGTCAGAAATATGATTCATGCTGAAGAGATC 2648
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 2649 ATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAGATTCATTTGAAGGAGTTGAA 2708
QY 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 2709 TATATGGCAGAGCTGGATGACCTCTGAGGAAATATGCTGGTCCATCTCTACTAGAT 2768
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 2769 CCTCCAGACTCGGCTACAGATAGTGTGATCTACCTGAATATTTATCCAGTAGAA 2828
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 2829 GATGATGTTATGGAAGGACAGACTCATTTGAGTCAGTGGCTGCTGCTGAGCGCCACTA 2888
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 2889 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 2948
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 2949 CCCCAAGTCAAGAAAGAGAAATTCAGTTATTTTTCCTCTGAATGCAAAACAGGTTTC 3008
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 3009 CGTCAATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAACGATCCAGTGT 3068
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 3069 GGGCTGCCTGCTCCCAAGTGATTTCAAGTGTCTATCAAGAGAGATAGCAATTCACAGT 3128
QY 501 GlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 3129 GGTGAATGGGAAGTCTTTGGCGCGCATGGATCTAATATCCAGTTGATGAGTCAGAGG 3188
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 3189 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCACT 3248
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 3249 TAGCTAAATCTCTGGAGAGTGCACAAGCGTCACTGACCGTGGTACTCACATTTCTGCTGC 3308
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 3309 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCCAGAGAATCCACACTGT 3368
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 3369 GTGTCCCTTTACAGCTATCAAGTCTTGAAGATGACCCCACTTGCACCAAAAGGAATTT 3428
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 3429 TGGGCCACCATTTTGGATTCAGCAGTCTCTTCTGACTATATCTACTCTCCAGAAATTTTC 3488
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 3489 TCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCTCTCATGATCTA 3548

QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 3549 CAGCCTGGAAAGAAATATCTACTGTCTGTTCATATATGTTGCTCTCAGGTCACTTG 3608
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 3609 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGTTGAATACCTTAGCCTCTCTAGGT 3668
QY 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 3669 TATGTGGTGTGATGATAGACCAAGGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGC 3728
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 3729 GCCTTTAAATATAAATGGTCAATAGAAATGACGATCAGGTGGAGGACTCCATAT 3788
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 3789 CTAGCTTCTCGATATGATTTTCACTTAGATCGTGGGCATCCACGGCTGCTCTAT 3848
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 3849 GGAGGATACCTCTCCCTGATGCAATTAATGCAGAGTTCAGATATCTTCAGGGTTGCTATT 3908
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 3909 GCTGGGGCCCACTGCTCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATG 3968
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 3969 GGTCAACCTCGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCGAA 4028
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 4029 AGTTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAAATGTC 4088
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 4089 CATTTTGCATACACCATATATTTACTGAGTTTTTTAGTGGGCTGGAAAGCCATATGAT 4148
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 4149 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTAT 4208
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 4209 GAATGCATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGTATTTGCTGCTCTAAAA 4268
QY 881 ValIle 882
Db 4269 GTGATA 4274

RESULT 11

AR631283 LOCUS 4829 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 12 from patent US 6844180.
ACCESSION AR631283
VERSION AR631283.1 GI:59770927
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4829)
AUTHORS Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 12 18-JAN-2005;
Ferring BV;;
NLM; Location/Qualifiers
source 1. 4829
/organism="unknown"
/mol_type="genomic DNA"

Db 2194 TGGTGAATATCGGTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAG 2253
Qy 680 lyTyrrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLeuPheGluG 700
Db 2254 GTTATGTGGTGTAGTATAGACACACAGGGATCTGTCACCGAGGGCTTAAATTTGAAG 2313
Qy 700 lyAlaPheLyTyrrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnT 720
Db 2314 GCGCTTTAAATATAAATATGGGTCAATAGAAATTTGACGATCAGTGGAGGACTTCAT 2373
Qy 720 YrLeuAlaSerArgTyrrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerT 740
Db 2374 ATCTAGCTTCTCATATGATTTTCATTTGACTTAGATCGTGTGGCATCCACGGCTGGTCT 2433
Qy 740 YrGlyGlyTyrrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 2434 ATGGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGGTTCCTA 2493
Qy 760 leAlaGlyAlaProValThrLeuTrpIlePheTyrrAspThrGlyTyrrThrGluArgTyrrM 780
Db 2494 TTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATA 2553
Qy 780 etGlyHisProAspGlnAnGluGlnGlyTyrrTyrrLeuGlySerValAlaMetGlnAlaG 800
Db 2554 TGGGTCAACCTTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAG 2613
Qy 800 lylLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
Db 2614 AAAAGTTCCCTCTGAACCAAAATCGTTTACTCTCTTACATGGGTTTCTGGATGAGANTG 2673
Qy 820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrrA 840
Db 2674 TCCATTTTGACATACACGATATATTACTGAGTTTTTTAGTGAGGGCTGGAAGCCATATG 2733
Qy 840 splLeuGlnIleTyrrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisT 860
Db 2734 ATTTACAGATCTATCTCAGGAGACACAGCATTAAGATTTCTGNAATCGGAGAACATT 2793
Qy 860 YrGluLeuHisLeuLeuHisTyrrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuL 880
Db 2794 ATGAACGTGATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTCGTCTCTAA 2853
Qy 880 ysValIle 882
Db 2854 AAGTGATA 2861

RESULT 12
AX608735 4829 bp DNA linear PAT 17-FEB-2003
LOCUS Sequence 12 from Patent WO0231134.
DEFINITION AX608735
ACCESSION AX608735
VERSION AX608735.1 GI:28404303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
Novel serine protease genes related to dppiv
Patent: WO 0231134-A 12 18-APR-2002;
Ferring BV (NL)
Location/Qualifiers
1. .4829
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 4829
Score: 4680.00 Matches: 882

Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 99.6% Indels: 2
DB: 6 Gaps: 0
US-10-825-632-1 (1-882) x AX608735 (1-4829)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTCGGACTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrrValGluArgTyrr 40
Db 274 GAGGAGATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrrHisGlyTyrrMetMet 60
Db 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATATCCAGAAATATATCATGGTCATCATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCGCACCATGATTTTCATTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
Qy 81 AspArgIleTyrrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrrSer 100
Db 454 GACAGAAATCTATTACCTTGCCATGTCTGTGAGAACAGAGAAATATACATGTTTTATCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCCAAACTATCATATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCCTCTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrrGlyMetTyrrSerArgGluGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGCAACACTGGACTATGGAATGTATTCTCCGAGAAGAAGAACTATTAA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrrAspTyrrHisGlnGlySerGly 160
Db 634 GAAAGAAAACGCGATTGGAACAGTCGGAATTTGCTTACGATTATATCACCAGGAAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCATCGTAAAGAGATGGAGGGCCCA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCCAACATACG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAAATTTATGCCCTCTGCTATCCAGCTGGATGCTTTTATACATAGCACGATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrrValHisAsnGluLeu 240
Db 874 TGGATATCTAATCATCTTAACAGAGAGAGAGAGACTCACTTATGTGCACAATAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTTCTCCAAGAA 993
Qy 261 PheAspArgTyrrSerGlyTyrrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACACTCCCAAGTGG 1053
Qy 281 LysIleLeuArgIleLeuTyrrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 1054 AAAATTTCTAGAAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATC 1113
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrrProLysThrGlyThr 320
Db 1114 ACATCCCTTATGTTGAAAACAGAGGGCAGATTCAITCCCGTTATCTTAAACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340

Db 1174 GCAATCCTAAAGTCACTTTTAAGATGTGAGAAATAATGATTGATGCTGAGGAAGGATC 1233
Qy IleAspValIleAspIleAspIleGlnProPheGluIleLeuPheGluIleValGlu 360
Db 1234 ATAGATGTCATAGTAAGGAACCTAAATCAACCTTTTGAGATTCTATTATTGAAGGATTGAA 1293
Qy TyrlleAlaArgAlaGlyTTPThrProGluGlyIleValAspIleLeuLeuAsp 380
Db 1294 TATATTGCCAGAGCTGGATGACTCTCTGAGGGGAAATATGCTTGGTCCACTACTAGAT 1353
Qy ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCAGACTCCCTACAGATAGTGTGATCTCACCTGATTTATTTATCCAGTAGAA 1413
Qy AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAAGCAGACACTATGAGTCAGTCCCTGATCTCTGACGCCACTA 1473
Qy IleIleTyrgluGluThrThrAspIleTTPIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAAACAAAGACATCTGATATAATATCCATGATCTTTTCACTGTTTT 1533
Qy ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysIleThrGlyPhe 460
Db 1534 CCCCAAGATCAGGAAGGAAATGAGTTTATTTTGCCTCTGAATGCAAAACAGGTTTC 1593
Qy ArgHisLeuTyrlleThrSerIleLeuLeuGluSerIleTyrlleArgSerSerGly 480
Db 1594 CGTCATTTATACAAATATACATCTATTTTAAAGAAAGCAATAATAACGATCCAGTGGT 1653
Qy GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGCTCGCTCTCCAAAGTGATTTTCAAGTGCTCTATCAAGAGAGGAGATAGCAATACCAAGT 1713
Qy GlyGluTTPGluValLeuGlyValHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTCTTTGGCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG 1773
Qy LeuValTyrlleGluValThrLysAspSerProLeuGluHisLeuTyrlleValValSer 540
Db 1774 CTGGTATATTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAAGT 1833
Qy TyrlleValProGluValThrArgLeuThrAspArgGlyTyrlleSerHisSerCysCys 560
Db 1834 TAGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACATTTCTTGCTGC 1893
Qy IleSerGluHisCysAspPhePheIleSerIleTyrlleSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCACTGTGACTCTTTATAGTAAGTATAGTAACTAACCAAGAAATCCACACTGT 1953
Qy ValSerLeuTyrlleLysSerSerProGluAspAspProThrCysIleThrLysGluPhe 600
Db 1954 GTGTCCCTTTTCAAGCTATCAAGTCTGGAAGTACCACTGACCTGCTGCTACTACATTTCT 2013
Qy TrpAlaThrIleLeuAspSetAlaGlyProLeuProAspTyrlleThrProGluIlePhe 620
Db 2014 TGGGCGCACCATTTTGGATTACGACAGTCTCTCTGACTATATCTCTCAGAAATTTTC 2073
Qy SerPheGluSerThrThrGlyPheThrLeuTyrlleMetLeuTyrlleProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTATACATTTGATGGGATGCTCTACAGGCTCATGATCTA 2133
Qy GlnProGlyLysIleTyrlleProThrValLeuPheIleTyrlleGlyGlyProGlnValGln 660
Db 2134 CAGCCTGGAAAGAAATATCTACTGTCTGTGTTTATATATATGTTGCTCTCTCAGGTGCACT 2193
Qy euValAsnAsnArgPheLysGlyValLysTyrllePheArgLeuAsnThrLeuAlaSerLeu 680
Db 2194 TGGTGAATATTCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAG 2253
Qy lYTyrlleValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
Db 2254 GTTATGTTGGTTGATGATAGACAAAGGGGATCTCTGTCAAGGAGGCTTAAATTTGAAG 2313

Qy 700 lYAlaPheIleTyrlleMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnT 720
Db 2314 CGGCTTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAT 2373
Qy 720 YrLeuAlaSerArgTyrlleAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerT 740
Db 2374 ATCTAGCTTCTCGATATGATTTTCACTTGAATCTGTGGCATCTGTGGCATCCACGCTGGTCT 2433
Qy 740 YrGlyGlyTyrlleSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 2434 ATGAGGATACCTCTCCCTGATGGCATTAATCAGAGGTGAGATATCTCAGGTTGCTA 2493
Qy 760 leAlaGlyValProValThrLeuTTPIlePheTyrlleAspThrGlyTyrlleThrGluArgTyrl 780
Db 2494 TTGCTGGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGATACAGGACGTATATA 2553
Qy 780 etGlyHisProAspGlnAsnGluGlnGlyTyrlleTyrlleGlySerValAlaMetGlnAlaG 800
Db 2554 TGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAG 2613
Qy 800 luLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
Db 2614 AAAAGTTCCCTCTGAACCAAAATCGTTTACTCTCTTACATCGTTTCTCGGATGAGAATG 2673
Qy 820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrl 840
Db 2674 TCCATTTTGACATACCAAGTATATTACTGATGTTTTTTAGTGAGGCTGGAAAGCCATATG 2733
Qy 840 pLeuGlnIleTyrlleProGlnGluArgHiserIleArgValProGluSerGlyGluHisT 860
Db 2734 ATTACAGATCTATCTCTCAGAGAGACACAGCATAAAGATTTCTGAAATCGGAGAACAT 2793
Qy 860 YrGluLeuHisLeuLeuHisTyrlleGlnGluAsnLeuGlySerArgIleAlaAlaLeuL 880
Db 2794 ATGAATCTGATCTTTTGACATCTACCTTCAGAAANACCTTGGATCAGTATTGTGCTCTAA 2853
Qy 880 ysValle 882
Db 2854 AAGTGATA 2861
RESULT 13
BC043124
LOCUS
DEFINITION
IMAGE:6410075), complete cds.
ACCESSION
BC043124.1 GI:27695449
VERSION
BC043124.1
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 4799)
AUTHORS
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klauer R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Mesny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez J., Helton E., Madan A., Young A.C., Shevchenko Y., Bouffard G.O., Baskiesley R.W., Touchman J.W., Green E.D., Bucken M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzyzanski M.I., Skalska U., Smaluk D.E., Schnerch A., Schein J.E., Jones S.J. and Marra M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
REFERENCE
2 (bases 1 to 4799)
Strausberg, R.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldio, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Kepel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, I., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27229207.
Location/Qualifiers
1. 4799
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:58057 IMAGE:6410075"
/tissue_type="Brain, enriched mouse brain 12.5dp"
/clone_lib="NIH BMAP_F00"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
1. 4799
/gene="Dpp8"
/db_xref="GeneID:74388"
/db_xref="MGI:1921638"
288. 2966
/gene="Dpp8"
/codon_start=1
/product="dipeptidylpeptidase 8"
/protein_id="AAH43124.1"
/db_xref="GI:27695450"
/db_xref="GeneID:74388"
/db_xref="MGI:1921638"
translation="WKIPSGRCNMAAAMETEOLGVIFETAECEGNGSGSDRPKLEP FVVERYSQQLKLLADTRKYHGMMAKAPDPMFKVTRDPPGPHSDRVVYLAAMSGEN RENTLFYSIEPTINRAVLMWSKPLLDLPQATLDYGMYSREELPERKIGTVGI AAYDHPGSGTFLFOAGSGIYHIXDGGPHGTQOPLRNLVETSCENIRMDPKLCPAD PDWIAFHSNDIWI SNLVTRERITTYVHNLAMNEEDPRSAVATFVLVEEFDYRSG YWPCQASRTSGSKILRIIYEEENDESEVEI IHVSPMLETRADSPFRYPKGTANPK VTFKMSIWDVAAGIIDVIDKELVQPFELIFEGVEYIARAGWTPGKHAWSLILDRS OTHLOI VLI SPFLPI PVEDDAMDQRILIESVPDSVTPLI IYEBTDTWINIHIDPHV FSGHEVELEPI FASECKTGPHLYKITSLIKESKYKRS SGLPASPDPKPIKESITI TSGHEVEIRHGSNIWDDEARKGVIFETSGDSPELHLVYTSVYANPGEVRLTRGYS HSCSLRSCDFFI SKYNSQNPCHVCVSLYKLSPPEDDPVHKTEFWATILDSAGPLPDY TPPEIFSPESDTFTGLYGLMKPHDLQGGKPYTFLFYGGPOVQLVNNRPFKGVKIFR LNTLASLGVVWVINDRGSCHRLKFGKAFKMGQIEIDDOVEGLQYLASQYDFIDL DRVGHGHSYGGYLSLMALMORSI FRYVATAGAPVTLWIFDYDTGYTERVMGHDPONEQ GYLGSGVAMQAEKFPSPENRLLHLHGLDENVNFHTSILLSPFLVRAGKPYDLQIYPO ERHSIRVPSGHEVHELHLHYLQENLGSRLAALKVI"

FEATURES
source
gene
CDS
ORIGIN
Alignment Scores:

Pred. No.: 0 Length: 4799
Score: 4536.50 Matches: 846
Percent Similarity: 97.8% Conservative: 18
Best Local Similarity: 95.8% Mismatches: 18
Query Match: 96.5% Indels: 1
DB: 9 Gaps: 1
US-10-825-632-1 (1-882) x BC043124 (1-4799)
Qy 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
Db 315 ATGGCAGCAGCAATGGAAACAGAA CAGCTGGGTGTCGAGATATTGTGAACCTGCAGAGTGT 374
Qy 21 GluGlu---AsnIleGluSerGlnAspArgProIysLeuGluProPheTyrValGluArg 39
Db 375 GAGGAGGGAATGGTGAATCA CAGGACCGCCCAACTGGAGCCCTTTTATGTGCGAACCG 434
Qy 40 TyrSerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMet 59
Db 435 TATCTTGAGTCAGCTGAATAAGCTGCTGCTGATACAGAAATACCAAGGCTACATG 494
Qy 60 MetAlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHis 79
Db 495 ATGGCTAAGCGCCACATGACTTTATGTTTGTAAAGAGGACGATCCAGATGGCCTCAC 554
Qy 80 SerAspArgIleTyrTyrIleuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 99
Db 555 TCAGACAGGGTCTATTACCTTGCCATGTCGTGGTGAACACAGAGAAATACACATATTTTAT 614
Qy 100 SerGluIleProIysThrIleAsnArgAlaValLeuMetLeuSerTyrLysProLeu 119
Db 615 TCTGAATCCCTTAAACCATCAACAGCAGCAGCTTAATGCTTTCTGGAAGCCCTT 674
Qy 120 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeu 139
Db 675 TTGGATCTTTTCAGCGGCACACTAGACTATGGATGTATTCTCGAGAGGAGAGTACTC 734
Qy 140 ArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySer 159
Db 735 AGAGAAGAAAGCGCATTTGGAACCGTGGGATCGCAGCTTATGATTACCAAGGAGGAGC 794
Qy 160 GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGln 179
Db 795 GGAACATCTCTGTTTCAAGCTGGCAGCGCATTTACCATATCAAGAGATGAGGAGCCACAT 854
Qy 180 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 199
Db 855 GGATTTACACACAGCCTTTGCGGCCCAATTTAGTGGAACTAGTTGTCCCATATACGA 914
Qy 200 MetAspProIysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAsp 219
Db 915 ATGGATCCAAATATTATGCTCTGCTGATCCAGACTGGATAGCTTTTCATTCACAGCAATGAT 974
Qy 220 IleTrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGlu 239
Db 975 ATTTGGATATCAACCTCTTAACAGGAGAGGAGGAGGATACATACGTACACAAATGAG 1034
Qy 240 LeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 259
Db 1035 CTAGCCAAACATGGAAGAGATCCAGATCAGCTGGTGTGTCACCTTTGTTCTTCAAGAA 1094
Qy 260 GluPheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGly 279
Db 1095 GAATTTGACAGATACTCTGCTACTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1154
Qy 280 GlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHis 299
Db 1155 GGTAAATCTTTAGAAATCTCTATGAGAAATGATGAATCTGAGGTGGAGATTTATTCAT 1214
Qy 300 ValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLeuThrGly 319
Db 1215 GTTACGTCCCCCATGTTGGAAACAGGAGGCGAGATTCCTTTCTGTTATCCCAAAACAGGC 1274

QY	320	ThrAlaAsnProIyValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArg	339	Db	2355	GTTTATGTGGTGTGGTGATAGACAAACAGGGGATCTCTGTACCGAGGACTTAAATTTTGA	2414
Db	1275	ACAGCAACCCAAAGGTCATCTTCAAGATGTCGAGATTTGTGTGATGCTCCAGGAGGG	1334	QY	700	GlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln	719
QY	340	IleIleAspValIleAspLysGluLeuIleGlnProPheGluIlePheGluGlyVal	359	Db	2415	GGCGCTTTAATATATAAATGGGTCAAAATAGAAATCGATGATCAAGTGAAGGACTCCAG	2474
Db	1335	ATTATAGATGTCATAGATAAAGAACTGGTTCAACCTTTCCGAGATCTCTGTGTGAGGAGTT	1394	QY	720	TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPsr	739
QY	360	GluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuLeu	379	Db	2475	TACCTAGCATCTCAGTATGACTTCATTTGATCGATCGATGGGATCCACGGCTGGTCC	2534
Db	1395	GAATATATTCAGAGCCGATGGACTCCAGAGGGAAACATGCTCGTGGTCATCTACTA	1454	QY	740	TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	759
QY	380	AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal	399	Db	2535	TATGTGGCTACCTCTCCGTATGGCAATTAATCGAGGTCCGATATCTTCGGGTGGCT	2594
Db	1455	GACCGTTCGCCAGACTCACTGCAGATAGTTCTGATCTCCCTGAGTATTATATCCAGTA	1514	QY	760	IleAlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyr	779
QY	400	GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro	419	Db	2595	ATTGCTGGGGCCCCAGTACCCCTGTGGATCTTCTATGATACAGATACAGGAGCCCTAT	2654
Db	1515	GAAGATGATGTCATGGACAGACAGACTTATAGATCAGTTCCTGACTCTGTGACACCA	1574	QY	780	MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla	799
QY	420	LeuIleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisVal	439	Db	2655	ATGGGTCACTCCAGCAAGATGAACAGGGCTACTTACTTAGGATCTGTGGCCATGCAAGCG	2714
Db	1575	CTGATCATCTATGAAGAAACACACAGACATCTGGATAAATATATCCAGATATTTTTCATGTT	1634	QY	800	GluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsn	819
QY	440	PheProGlnSerHisGluGluLeuGluPheIlePheAlaSerGluCysLysThrGly	459	Db	2715	GAGAAAGTTCCCTCAGAACCAACCGGTACTCTCTTCATCGGTCTTGGATGAGAAT	2774
Db	1635	TTTCTCAACTCATGAAGATGAATTTGAGTTTATTTTTCCTCTGAAATGCAAAACAGGT	1694	QY	820	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr	839
QY	460	PheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSer	479	Db	2775	GTTTCACTTTGACACACACAGTATATTGCTGAGTTTTTTAGTGAGGGCTGGAAGGCCATAT	2834
Db	1695	TTTCTGTCATCTGTATTAATTAACATCCATTTTAAAGGAGGACCAATATAAAGCGTCCAGT	1754	QY	840	AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis	859
QY	480	GlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluLeuIleAlaIleThr	499	Db	2835	GACTTACAGATCTATCTCTCAGGAGGACACAGCATCAGAGTTCTGAGTCTGGAGAACAC	2894
Db	1755	GGTGACATCTGCTCCCAAGGATTTCAAGTGCTTATTAAGAGGAGGACCAATATAAAGTTACC	1814	QY	860	TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLeu	879
QY	500	SerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	519	Db	2895	TATGAACTGCACCTGCTCCACTACTTCAGGAGAACCTTCGATCGGTATCGTCTGCTG	2954
Db	1815	AGTGTGTAATGGGAAGTACTTGGCCGCGCATGGATCTAATATCTGGGTGATGAGCCAGA	1874	QY	880	LysValIle	882
QY	520	ArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValVal	539	Db	2955	AAAGTGATA	2963
Db	1875	AAGCTGGTGATCTTTGAAGGACCAAAAGACTCTCTTTGGAGCATCACCTGTACGTGACC	1934	RESULT 14			
QY	540	SerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys	559	BC059222			
Db	1935	AGTTATGCAAAACCTTGGAGAAAGTGTGAGGCTGACTGACCGTGGCTACTCACACTCCCTGC	1994	LOCUS			
QY	560	CysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHis	579	DEFINITION			
Db	1995	TGCCTCAGCCGCGCATGTGTGACTTCTTATAGTAAGTACAGCAACCAAGAAATCCACAC	2054	IMAGE:6410075), complete cds.			
QY	580	CysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu	599	ACCESSION			
Db	2055	TGTGTGTCCTCTACAAACTCTCAAGTCTTGAGGATGACCCAGTTCTATTAACAAAGGAA	2114	BC059222.1			
QY	600	PheTyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIle	619	VERSION			
Db	2115	TTTGTGGCCACCAATTTGGATTACGAGGTCCTCTCTGACTACACCCCTCCAGAAAT	2174	KEYWORDS			
QY	620	PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAsp	639	SOURCE			
Db	2175	TTTTCTTTTGAAGTACTACTGGAATTTACACTGTATGGAATTTGTATAGCTCATGAC	2234	Mus musculus (house mouse)			
QY	640	LeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGln	659	Mus musculus			
Db	2235	CTACACCTCTGGAAGAAATACCCACCTGTGTATTATATATGTTGTTGCTCCAGGTGACG	2294	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
QY	660	LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAspLeu	679	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;			
Db	2295	CTGGTGAAACATTCGGTTTAAAGGAGTCAAGTATTTCCGCTGAACACCTTGGCTCCCTG	2354	Sciurognathi; Muridae; Murinae; Mus.			
QY	680	GlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLysPheGlu	699	1 (bases 1 to 4799)			
				REFERENCE			
				AUTHORS			
				Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
				Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,			
				Altshuler, F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
				Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,			
				Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,			
				Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,			
				Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,			
				Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,			
				Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,			
				McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,			
				Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,			
				Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
				Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,			
				Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
				Bouffard, G., Blakeley, R.W., Touchman, J.W., Green, E.D.,			
				Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			

QY	320	ThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArg	339	Db	2355	GGTTATGTGGTGTGGTGATAGACACAGGGGATCTGTCCAGGAGACTTAAATTTGAA	2414
Db	1275	ACAGCAACCAAGAGTCACTTTCAAGATGTCGAGAGATTGTTGTGATGTCGAGGAGG	1334	QY	700	GlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln	719
QY	340	IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVal	359	Db	2415	GGCCCTTTAAATATAAATGGGTCAAAATAGAAATCGATGATCAAGTGGAGGACTCCAG	2474
Db	1335	ATTATAGATGTCATAGATAAAGAACTGGTTCAACCTTTTCGAGATTCCTGTTTGGAGGAGTT	1394	QY	720	TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSer	739
QY	360	GluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeu	379	Db	2475	TACCTAGCATCTCAGTATGATCTTCATTTGACTTGGATCGAGTGGGCATCCACGGCTGGTCC	2534
Db	1395	GAATATATGCCAGAGCCGATGGACTCCAGAGGGAACAACTGCTGCTCCACTACTA	1454	QY	740	TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	759
QY	380	AspArgSerGlnThrArgIleGlnIleValLeuIleSerProGluLeuPheIleProVal	399	Db	2535	TATGGTGGCTACCTCTCCCTGATGGCATTAAATGACAGGTCGGATATCTTCCGGGTGGCT	2594
Db	1455	GACCGTTCCAGACTCACCTGCAGATGTTCTGATCTCCCTGAGTATTATCATCCAGTA	1514	QY	760	IleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyr	779
QY	400	GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro	419	Db	2595	ATTGCTGGGGCCCCAGTCACCTGTGGATCTTCTATGATACAGGATACACGGAGCGCTAT	2654
Db	1515	GAAGATGATGCCATGGACAGACAGACTTATAGAGTCAGTTCCTGACTCTGTGACCCA	1574	QY	780	MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla	799
QY	420	LeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisVal	439	Db	2655	ATGGGTCACTGACCAAGAAATGAACAGGGCTACTACTTAGGATCTGTGGCCATGCAAGCG	2714
Db	1575	CTGATCATCTATGAAGAAACAACAGACATCTGGATAAATATCCAGATATTTTTCATGTT	1634	QY	800	GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn	819
QY	440	PheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGly	459	Db	2715	GGAAGATTCCCTCAGAACCAACCGGTTACTCTCTTGCATGGGTCTTGTGATGAGAT	2774
Db	1635	TTTCTCAAACTCATGAAGATGAAATTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGT	1694	QY	820	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr	839
QY	460	PheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSer	479	Db	2775	GTTCATTTGCACACACAGTATATGCTGAGTTTTTTAGTGAGGCTGGAAGACCATAT	2834
Db	1695	TTTCTCATCTGTATAAATACATCCATCTTTAAAGGAGAGCAAAATATAAACCGTCCAGT	1754	QY	840	AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis	859
QY	480	GlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThr	499	Db	2835	GACTTACAGATCTATCTCTCAGGAGGACAGCATCAGAGTTCTCGAGTCTGGAGAACAC	2894
Db	1755	GGTGACTACTCTGCCCAAGTATTTCAAGTGTCTTATCAAGAGAAATAACAATTACC	1814	QY	860	TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeu	879
QY	500	SerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	519	Db	2895	TATGAATGCACTGCTGCTCCACTTACCTTACAGGAACCTTGGATCGGATCGCTGCTCTG	2954
Db	1815	AGTGTGAATGGGAAGTACTTGGCCGCATGGATCTAATATCTGGGTTGATGAGCCAGA	1874	QY	880	LysValIle	882
QY	520	ArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLysTyrValVal	539	Db	2955	AAAGTGATA	2963
Db	1875	AAGCTGGGTACTTTGAAGGCACCAAGACTCTCTTTGGAGCATCCTCTGACTGTACC	1934	RESULT 15			
QY	540	SerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys	559	AR448400		2797 bp	DNA
Db	1935	AGTTATGCAAAACCCCTGGAGAGAGTGGTGAGGCTGACTGACCGTGGCTACTCACACTCTGC	1994	LOCUS	AR448400	1103 from patent	US 6673549.
QY	560	CysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHis	579	DEFINITION	AR448400		
Db	1995	TGCTTCAGCGGCATTGTGACTTCTTCAATAAGTAAGTACAGCAACCAAGAAATCCACAC	2054	ACCESSION	AR448400.1	GI:42676724	
QY	580	CysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu	599	VERSION			
Db	2055	TGTGTGTCCTCTACAAACTCTCAAGTCTTGAGGATGACCCAGTTCTATAAAACAAAGGAA	2114	KEYWORDS			
QY	600	PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIle	619	SOURCE			
Db	2115	TTTGTGGCCACCAATTTGGATTGACAGAGTCTCTCTGACTACACCCCTCCAGAAATT	2174	ORGANISM			
QY	620	PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAsp	639	REFERENCE			
Db	2175	TTTCTCTTGAAGTACTACTGGATTATACATGTATGGAATGTTGATTAAGCTCATGAC	2234	1 (bases 1 to 2797)			
QY	640	LeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGln	659	Furness,L.M. and Buchbinder,J.L.			
Db	2235	CTACAACTCGGAAAGAAATACCCCACTGTGTATTATATATATGTTGTCCTCCAGGTGAC	2294	Genes expressed in C3A liver cell			
QY	660	LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeu	679	Patent: US 6673549-A 1103 06-JAN-2004;			
Db	2295	CTGGTGAACAATCGGTTAAAGAGAGTCAAGTATTTCCGCTCGAACCCCTGGCTCCCTG	2354	Incyte Corporation; Palo Alto, CA			
QY	680	GlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu	699	Location/Qualifiers			
				1. .2797			
				/organism="unknown"			
				/mol_type="genomic			
				DNA			
				ORIGIN			
				Alignment Scores:			
				Pred. No.:	0	Length:	2797
				Score:	4443.50	Matches:	840
				Percent Similarity:	99.5%	Conservative:	0
				Best Local Similarity:	99.5%	Mismatches:	1
				Query Match:	94.5%	Indels:	3
				DB:	6	Gaps:	1
				US-10-825-632-1 (1-882) x AR448400 (1-2797)			

QY 42 TrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMetAla 61
Db 3 TGGAGTCAGCTTAAAGAGCTGCTTGCAGATACAGAGAAATATCATGGCTACATGATGGCT 62
QY 62 LysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSerAsp 81
Db 63 AAGGCACACATGATATTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAGAC 122
QY 82 ArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSerGlu 101
Db 123 AGAATCTATTACCTTGCATGCTGCTGTGAGAACAGAGAAATATACACTGTTTATTCTGAA 182
QY 102 IleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeuAsp 121
Db 183 ATTCCCAAAATCATCAATAGACAGCAGCTCTTAATGCTCTCTTGAAGCCTCTTTTGGAT 242
QY 122 LeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuLeuArgGlu 141
Db 243 CTTTTTTCAGGCAACACTGGACTATGGAATGTATTCGAGAGAAAGAACTATTAGAGAA 302
QY 142 ArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGlyThr 161
Db 303 AGAAACCAATGGACAGTCGGAATGCTTCTTACGATTTATCCACAGAGAAAGTGAACA 362
QY 162 PheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGlyPhe 181
Db 363 TTTCTGTTTCAAGCCGCTAGTGGAAATTTATCCGTAAGATGGAGGCCACAGGATTT 422
QY 182 ThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAsp 201
Db 423 AGCCAAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATGGAT 482
QY 202 ProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIleTrp 221
Db 483 CCAAAATATGCGCTGCTGATCCAGACTGGATTCCTTTATACATAGAACAGATATTGG 542
QY 222 IleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeuAla 241
Db 543 ATATCTAATCATCGTAACAGAGAGAAAGAGGACTCATTTATGTGCACAAATGAGTAGCC 602
QY 242 AsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPhe 261
Db 603 AACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTCTTCCAAAGAAATTT 662
QY 262 AspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLys 281
Db 663 GATAGATATTCGGCTATTGGTGTGTCGAGTCCAAAGCTGAAACAACTCCCGAGTGGTAAA 722
QY 282 IleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThr 301
Db 723 ATCTTAGAATTTATATGAGAAATGATGAATCTGAGTGGAAATTTATCATGTACA 782
QY 302 SerProMetLeuGluThrArg-ArgAlaAspSerPheArgTyrProLysThrGlyThrAl 321
Db 783 TCCCTATGTTGGAAACCAAGCCAGGCAGGATTCATTCGTTATCCTTAAACAGGATACAGC 842
QY 321 aAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleI 341
Db 843 AAATCCTTAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGGAAGGAGGATCAT 902
QY 341 eAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyr 361
Db 903 AGATGTCATAGATAGGAACATAATCAACCTTTTGAATTTCTATTGAGAGGATGGAATA 962
QY 361 rIleAlaArgAlaGlyTyrThrProGluGlyLys- - - -TyrAlaTrpSerIleLeuLeuA 380
Db 963 TATTGCCAGAGCTGGATGGACTCCTGAGGGAATAGTGTGCTTGTGCTCCACTACTAG 1022
QY 380 sPArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValG 400
Db 1023 ATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTTATTATCCCAAGTAG 1082

QY 400 LuAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProL 420
Db 1083 AAGATGATGTTATGAAAGGAGAGACTCATTTAGTCAGTCCTGATTTCTGACGCCAC 1142
QY 420 euIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValP 440
Db 1143 TAAATATCTATGAAGAAACACAGACATCTGGATAAATATATCATGACATCTTTCAITGTTT 1202
QY 440 heProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyP 460
Db 1203 TTCCCAAGTCCAGAGAGGAAATTCAGTTATTTTGGCTCTGAAATCAAAACAGGTT 1262
QY 460 heArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerG 480
Db 1263 TCCGTCATTTATACAAAATTTATCTATTTTAAAGGAAGCAATATAAAGCATCCAGTG 1322
QY 480 lyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrS 500
Db 1323 GTGGCTGCTGCTCCAAAGTGTATTTCAAGTCTCTATCAAGAGAGAGATAGCAATTACCA 1382
QY 500 erGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgA 520
Db 1383 GTGGTCAATGGGAAGTCTTGGCCGATGGATCTAATATATCAAGTTGATGAAGTCAGAA 1442
QY 520 rgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValS 540
Db 1443 GCGTGATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACTGTAGGTAGTCA 1502
QY 540 erTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysC 560
Db 1503 GTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTGCT 1562
QY 560 ysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisC 580
Db 1563 GCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACAGAGAAATCCACACT 1622
QY 580 ysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluP 600
Db 1623 GTGTGTCCCTTTACAAAGCTATCAAGTCTCGAAGATGACCCCAACTTGCAGAAACAAAGGAAT 1682
QY 600 heTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIleP 620
Db 1683 TTTGGGCCCACTTTTGGATTCAGCAGGTCTCTTCTCTGACTACTACTCTCCAGAAATTT 1742
QY 620 heSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspL 640
Db 1743 TCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAACAGCCTCATGATC 1802
QY 640 euGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnL 660
Db 1803 TACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTTCATATATATGGTGGTCTCAGGTGCA 1862
QY 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
Db 1863 TGTGGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAAATACCTTAGCTCTCTAG 1922
QY 680 lyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
Db 1923 GTTATGTGTTGTAGTAGACACAGGGATCTCTGTCCAGGGCTCTAAATTTTCAAG 1982
QY 700 lyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnT 720
Db 1983 GCGCCTTTAAATATAAATGGGTCAATAGAAATGACGATCAGGTGGAAGGATCCCAAT 2042
QY 720 yrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerT 740
Db 2043 ATCTAGCTTCTCGATATGATTTTCTTGTAGTCTGATGCTGGGATCCAGGCTGGTCT 2102
QY 740 yrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 2103 ATGAGGATACCTCTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTCAGGGTCTCTA 2162
QY 760 leAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrM 780

Db	2163	TTGCTGGGGCCCCAGTCACCTCTGTGGATCTTTATGATACAGATACACGGAAACGTTATA	2222
Qy	780	etGlyHisProAspGluAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG	800
Db	2223	TGGGTACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCCAAGCAG	2282
Qy	800	luLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnV	820
Db	2283	AAAAAGTTCCCTCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCCTGGATGAGAAATG	2342
Qy	820	alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA	840
Db	2343	TCCATTTTGCACATACCAAGTATATTACTGAGTTTTTTTAGTGAGGGCTGGAAAGCCATATG	2402
Qy	840	epLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHist	860
Db	2403	ATTTACAGATCTATCTCTCAGGAGAGACACAGCATNAGAGTTCTGAAATCGGGAGAACATT	2462
Qy	860	yrGluLeuHisGluLeuHisTyrTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuL	880
Db	2463	ATGAACCTGCATCTTTTGCACTTACCTTCAGAAAACCTTGGATCAGGTATTGCTGCTCTAA	2522
Qy	880	ysValIle	882
Db	2523	AAGTGATA	2530

Search completed: April 15, 2006, 03:40:39
Job time : 11208 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 00:03:07 ; Search time 1249 Seconds
(without alignments)
4706.372 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAAAMETQLGVETPFTADC.....HLLHYIQENLGRSIAALKVI 882

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10825632/runat_14042006_092558_4645/app_query.fasta_1
-DB=N_Geneseq -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss08
-USER=US10825632 @CGN_1_1_727 @runat_14042006_092558_4645 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	100.0	2649	ABX12255	Abx12255 cDNA enco
2	4700	100.0	2649	14 ADV43981	Adv43981 Human psy
3	4700	100.0	2671	6 ABK83322	Abk83322 cDNA enco
4	4700	100.0	3106	6 ABK12892	Abk12892 Human pro

5	4700	100.0	3120	4 AAC85694	Aac85694 Nucleotid
6	4700	100.0	3120	6 AAD38956	Aad38956 Human dip
7	4700	100.0	3143	6 AAH99934	Aah99934 cDNA enco
8	4695	99.9	2643	6 AAH99935	Aah99935 Coding se
9	4680	99.6	4829	6 ABK83327	Abk83327 cDNA enco
10	4443.5	94.5	2797	12 ADL13374	Adl13374 Human ste
11	4397.5	93.6	2952	10 ACA92421	Aca92421 DNA enco
12	4385.5	93.3	4685	6 ABK83332	Abk83332 cDNA enco
13	4385	93.3	4676	6 ABK83331	Abk83331 cDNA enco
14	4353	92.6	2929	10 ACA92425	Aca92425 cDNA enco
15	4118	87.6	2349	14 ADV43982	Adv43982 Human psy
16	4118	87.6	2842	6 ABN59774	Abn59774 Novel hum
17	4092.5	87.1	4523	6 ABK83325	Abk83325 cDNA enco
18	3970.5	84.5	2510	6 AAD23843	Aad23843 Human pro
19	3901	83.0	2696	12 ADI16386	Adi16386 Human pro
20	3817	81.2	2702	13 ADT04072	Adt04072 Human pro
21	3771	80.2	2668	6 ABN59775	Abn59775 Novel hum
22	3661.5	77.9	4309	6 ABK83328	Abk83328 cDNA enco
23	3641	77.5	2251	10 ADE79035	Ade79035 Human pro
24	3597.5	76.5	2463	10 ACA92424	Aca92424 DNA enco
25	3364.5	71.6	2161	4 AAH15009	Aah15009 Human CD-
26	3087	65.7	2083	7 ADR41222	Adr41222 Human enco
27	2870	61.1	2617	6 ABQ75955	Abq75955 Human PMW
28	2870	61.1	3716	6 ABQ75955	Abq75955 Human PMW
29	2870	61.1	4147	12 ADQ8753	Adq8753 Human tum
30	2870	61.1	4147	13 ADQ84273	Adq84273 Human tum
31	2870	61.1	4147	13 ACN40423	Acn40423 Tumour-as
32	2870	61.1	4219	6 ABK83335	Abk83335 cDNA enco
33	2870	61.1	4302	6 ABK83333	Abk83333 cDNA enco
34	2870	61.1	4398	13 ADS10267	Ads10267 Human the
35	2863	60.9	3024	6 AAD38954	Aad38954 Human dip
36	2862	60.9	2660	6 ADI16689	Adi16689 Human NOV
37	2862	60.9	2660	6 ADI16687	Adi16687 Human dip
38	2835	60.3	2495	6 AAD38957	Aad38957 Human dip
39	2833	60.3	3287	6 AAD38955	Aad38955 Alternati
40	2832	60.3	3287	6 AAD38955	Aad38955 Alternati
41	2820.5	60.0	4180	12 ADN42341	Adn42341 Human cDN
42	2820.5	60.0	4263	6 ABK83339	Abk83339 cDNA enco
43	2820.5	60.0	4359	13 ADS10269	Ads10269 Human the
44	2801	59.6	1821	6 ABV76411	Abv76411 Dipeptidy
45	2763	58.8	2751	6 AAD38311	Aad38311 Murine di

ALIGNMENTS

RESULT 1
ABX12255
ID ABX12255 standard; cDNA; 2649 BP.
XX
AC ABX12255;
XX
DT 19-MAY-2003 (first entry)
XX
DE cDNA encoding human serine protease HIPHUM46.
XX
KW Human; ss; gene; HIPHUM46; serine protease; gene therapy; osteoarthritis;
KW serine protease activity modulation; dipeptidyl peptidase activity;
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
KW Alzheimer's disease; paraspranuclear palsy; Huntington's disease;
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
KW irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
KW multiple sclerosis; chromosome 15q21-q22.
XX
OS Homo sapiens.

Key	Location/Qualifiers
FT	1..2649
FT	/*tag= a
FT	/product= "HIPHUM46"
XX	/note= "Serine protease"

PN GB2374869-A.
XX 30-OCT-2002.
PD 22-JAN-2002; 2002GB-000011404.
XX 23-JAN-2001; 2001GB-000011760.
XX (GLAX) GLAXO GROUP LTD.
PA Edbrooke MR, Lewis AP;
PI WPI: 2003-150703/15.
DR P-PSDB; ABU07720.
XX Identifying modulators of serine protease activity useful for treating
PT musculoskeletal diseases, by contacting cell expressing a novel serine
PT protease polypeptide with a compound and monitoring serine protease
PT activity.
XX Claim 12; Page 22-26; 38pp; English.
XX The invention relates to a method of identifying a substance that
CC modulates serine protease activity, comprising contacting a cell such as
CC a neuronal cell, lung cell, intestinal cell or a cell infected with a
CC virus, expressing a serine protease polypeptide (HrPHUM 46), or its
CC variant having dipeptidyl peptidase activity, or a serine protease
CC isolated from the cell with a test substance and monitoring for serine
CC protease activity. The method is useful for identifying a substance that
CC modulates serine protease activity. A modulator of the serine protease is
CC useful in the manufacture of a medicament for treatment or prophylaxis of
CC a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus
CC infection, Alzheimer's disease, parasupranuclear palsy, myotonic
CC dystrophy, Huntington's disease or amyotrophic lateral sclerosis.
CC Additional disease that may be treated using modulators of the serine
CC protease include malabsorption syndromes, irritable bowel syndrome, lung
CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,
CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,
CC dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple
CC sclerosis. The present sequence represents cDNA encoding the human serine
CC protease HrPHUM46 which is located on chromosome 15q21-q22
XX Sequence 2649 BP; 803 A; 514 C; 585 G; 747 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 0 Length: 2649
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-825-632-1 (1-882) x ABX12255 (1-2649)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
Db 1 ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTGAGATATTGAAACTGGCGACTGT 60
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 61 GAGCAGAAATATTGAATACAGGATCGGCTAAATTTGGAGCGCTTTTATGTGTGACCGGTAT 120
Qy 41 SerTpsSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 121 TCCTGGAGTCAGCTTAAAAAGCTGCTTCGCATACAGAAATATCATGCTGCATCATGATG 180
Qy 61 AlaValAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 181 GCTAAGGCACCATCATGTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 240
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 241 GACAGAAATCTATTACCTTCCCATGCTGGTGAGACAGAGAAATAACACTGTTTATTCT 300

Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 301 GAAATTTCCAAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGAAGCCTCTTTTG 360
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 361 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTTCGAGAAGAAGACTATTAGA 420
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 421 GAAAGAAAACGCATTGGAAACAGTCGGAATTCCTTACGATTATCCACAGAGAGTGA 480
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 481 ACATTTCTGTTCGAGCGGTAGTGGAAATTTATCAGCTAAAGATGGAGGGCCACAGA 540
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 541 TTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACG 600
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 601 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTGCTTTATACATAGCAACGAT 660
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 661 TGGATATCTACATCGTAACCCAGAGAAAGAGGAGACTCACTTATGTGCACAATAGCT 720
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 721 GCACATCGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCCAAGAGAA 780
Qy 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrPheThrProSerGlyGly 280
Db 781 TTGTAGATATATTCGGCTATTTGGTGTGTCCTCAAGCTCAACCACTCCAGTGGTGGT 840
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 841 AAAATCTTTAGAAATCTATATGAGAAATGATGATCTGAGGTGGAAATATTATCATG 900
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 901 ACATCCCTTATGTTGGAAACCAAGGAGGCGAGATTCATTCGTTATCTTAAACAGGTACA 960
Qy 321 AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle 340
Db 961 GCAATCTCTAAAGTCACTTTTAAAGATGTGAGAAATAATGATTGATGCTGAAGGAAGATC 1020
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1021 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCTTATTTGAAGGAGT 1080
Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1081 TATATTGCCAGAGCTGGATGGACTCTCAGGGAAAAATATGCTTGGTCCATCTACTAGAT 1140
Qy 381 ArgSerGlnThrArgGluGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1141 CGCTCCAGACTCGCTGCAGATAGTGTGTGATCTCACCTGAATTTATTTATCCCAAGTAG 1200
Qy 401 AspAspValMetGluArgGluArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1201 GATGATGTTATGGAAGGCGAGACTCATTTGAGTCAGTGCCTGATTCTGTGACGCCACTA 1260
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1261 ATTATCTATGAAGAAACACACATCTGGATTAATATCCATGACATCTTTTCATGTTTT 1320
Qy 441 ProGlnSerHisGluGluGluLeuPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1321 CCCCAGAGTCACAGAGGAAATGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1380

Qy 461 ArgHisLeuTyrIleThrSerIleLeuLysGluSerIleTyrIleLysArgSerSerGly 480
Db 1381 CGTCAATTTATACAAAATACATCTATTTTAAAGGAAACAAATATAAACCATCCAGTGT 1440
Qy 481 GlyLeuProAlaSerAspPheLysCysProIleLysGluLleAlaIleThrSer 500
Db 1441 GGCTCCCTCCAGGTGATTTCAAGTGTCTTATCAAGAGGAGATAGCAATACCAT 1500
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1501 GGTGAATGGAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1560
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1561 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGT 1620
Qy 541 TyrValAsnProGlyValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1621 TACGTAAATCCCTGGAGAGTGACAGGCTGACTGACCTGGCTACTCACATTTCTTGCTGC 1680
Qy 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1681 ATCAGTCAGCAGCTGTGACTCTTTATAGTAGTATAGTAACCAAGAGATCCACACTGT 1740
Qy 581 ValSerLeuTyrIleLysLeuSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1741 GTGTCCCTTTACAGCATATCAAGTCTCAAGATGACCACTTGCAAAAACAAAGGAATTT 1800
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1801 TGGGCCACCATTTGGATTGACAGGCTCTTCTGACTATATCTCTCCAGAAATTTTC 1860
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1861 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTA 1920
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1921 CAGCCTGGAAGAAATATCTACTGTGTGTTTATATATATGTTGCTCCTCAGGTGCAAGTTG 1980
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1981 GTCAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATATACCTTAGCTCTCTAGGT 2040
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2041 TATGTGGTGTGTAGTGATACACACAGGGGATCCTGTCCACGAGGCTTAATTTGAAGGC 2100
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2101 GCCTTTAATATATAAATGGTCAATAGNAATTGACGATCAGGTGGAAGGACTCCAATAT 2160
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2161 CTAGCTTCGATATGATTTTCACTTACATCTGTGGCATCTCCAGGCTGGTCTCTAT 2220
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2221 GGAAGGATACCTCTCCCTGATGGATTAATGCAAGAGTCAAGATATCTTCAGGGTGTCTATT 2280
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2340
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2341 GGTCAACCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2400
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2401 AAGTTCCTCTCTCAACCAATCGTTTACTGTCTTTACATGTTGTTCTCTGGATGAGAAATGC 2460
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

Db 2461 CATTTTGCATACACAGTATATTTCTAGTATTTTTTAGTGAGGCTGGAAGCCATATGAT 2520
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2521 TTACAGATCTATCTCTCAGGAGACACAGACATAGAGTTCTCTGAATCTGGAGAACATTAT 2580
Qy 861 GluLeuHisLeuLeuHisTyrIleGlnGlnAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2581 GAACTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTATTGTCTGCTCTAAA 2640
Qy 881 ValIle 882
Db 2641 GTGATA 2646
RESULT 2
ADV43981
ID ADV43981 standard; cDNA; 2649 BP.
XX
AC ADV43981;
XX
DT 10-MAR-2005 (first entry)
XX
DE Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 1609.
XX
DE microarray; psychoneuroendocrinimmune; chronic fatigue;
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW cancer; neoplasm; infection; expressed sequence tag; ss.
XX
OS Homo sapiens.
XX
PN W02004108899-A2.
XX
PD 16-DEC-2004.
XX
PF 04-JUN-2004; 2004WO-US017686.
XX
PR 04-JUN-2003; 2003US-0475915P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nicholson A, Vernon SD;
XX
DR WPI; 2005-031682/03.
XX
PT New microarray comprising probes for genes involved in
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT condition associated with PNI activity, e.g., inflammatory or infectious
XX diseases.
XX
PS Claim 1; SEQ ID NO 1609; 254pp; English.
XX
CC The invention relates to a new microarray which comprises probes for
CC genes involved in psychoneuroendocrinimmune (PNI) activity. The
CC microarray is useful in diagnosing a condition associated with PNI
CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
CC cancer and infection. The present sequence represents a
CC psychoneuroendocrinimmune gene expressed sequence tag. Note the
CC specification mentions SEQ ID NO of up to 3314 but only sequences up to
CC SEQ ID NO 1829 are provided.
XX
SQ Sequence 2649 BP; 804 A; 514 C; 585 G; 746 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 2649
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0
US-10-825-632-1 (1-882) x ADV43981 (1-2649)

QY	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
DB	1	ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTGAGATATTGAAACTCGGACGTGT	60
QY	21	GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
DB	61	GAGGAGAATATTGAATCACAGGATCGCCCTAAATTTGGAGCCCTTTTATGTGTGAGCGGTAT	120
QY	41	SerTyrSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
DB	121	TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGGATACCAGAAAATATCATGCTACATGATG	180
QY	61	AlaLysAlaProHisAspPheMetPheValLysArgAenAspProAspGlyProHisSer	80
DB	181	GCTAAGGCCACCATGATTTCTGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA	240
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer	100
DB	241	GACAGATCTATTACCTTTCGATCTGTGTTGAGACAGAGAAATACACTGTTTTATTCT	300
QY	101	GluIleProLysThrIleAenArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu	120
DB	301	GAATATCCCAAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCCTCTTTG	360
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
DB	361	GATCTTTTTCAGGCCAACCTGGACTATGGAATGTATTTCTCGAAGAAGAACATTTAAG	420
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
DB	421	GAAGAAGAAACGATTTGGACAGCTGGGAATTTGCTTTTACGATTATCACCAAGAGATGGA	480
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
DB	481	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAAGGA	540
QY	181	PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet	200
DB	541	TTTACGCCAACACCTTTAAGSCCCAACTAGTGGAACTAGTTGTGCCAACATACCGATG	600
QY	201	AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle	220
DB	601	GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTATACATAGCAACGATATT	660
QY	221	TyrIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu	240
DB	661	TGATATCTAAATCATCGTAACAGAGAAAGAGAGACTCATTTATGTGCACAAATGAGCTA	720
QY	241	AlaAsnMetClnGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
DB	721	GCCACATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA	780
QY	261	PheAspArgTyrSerClnTyrTyrTyrCysProLysAlaGluThrThrProSerGlyGly	280
DB	781	TTTGATAGATATTCTGGCTATTGTTGTGTCTCCAAAGCTGAAACAACTCCCAAGTGGT	840
QY	281	LysIleLeuArgIleLeuTyrGlnGluAsnAspGluSerGluValGluIleIleHisVal	300
DB	841	AAAATTTCTTAGAATTTCTATACAGAAATATGATGATCTGAGGTGGAAATTTATTCATGTT	900
QY	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrClnThr	320
DB	901	ACATCCCTCATGTTGGAAACCAAGAGGGCAGATTCATTCCGTTATCTCTTAAACAGATACA	960
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlnClyArgIle	340
DB	961	GCAATCTTAAGTCACTTTTAAGATGTGAGAAATATGATTTGATGCTCAGGAAGGATC	1020
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluClyValGlu	360
DB	1021	ATAGATGTCATAGATAAGAACTAACTCAACCTTTTGAGATTTCTATTGAGGAGGTGAA	1080
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuLeuAsp	380

Db 2161 CTAGCTTCGATAGATTTCAATGACTTAGATCGTGTGGGATCCACGGCTGTCCTAT 2220

Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760

Db 2221 GGAGATACCTCTCCCTGATGGCATTAATGCAGAGTCAGATATCTTCAGGGTTCCTATT 2280

Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780

Db 2281 GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACAGGAACGTTATATG 2340

Qy 781 GlyHisProAspGlnAsnGlnGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlu 800

Db 2341 GGTCAACCTGCACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGAGAA 2400

Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820

Db 2401 AAGTTCCTCCCTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTC 2460

Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

Db 2461 CATTTTGCATACACAGTATATTACTAGTTTTTTAGTGAGGGCTGGAAGCCATATGAT 2520

Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860

Db 2521 TTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCCTGAATCGGGAGACATTAT 2580

Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880

Db 2581 GAACTGCACTCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGTATTGCTCTTAATA 2640

Qy 881 ValIle 882

Db 2641 GTGATA 2646

RESULT 3

ABK83322

ID ABK83322 standard; cDNA; 2671 BP.

XX AC ABK83322;

XX 12-AUG-2002 (first entry)

DT cDNA encoding human DPPIV related serine protease DPPP-1.

DE Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;

XX DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;

KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;

KW heart failure; hypertension; urinary retention; osteoporosis; cancer;

KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;

KW dyskinetia; reproductive disorder; inflammatory disorder;

XX metabolic disorder; gene; ss.

XX Homo sapiens.

OS WO200231134-A2.

XX PN 18-APR-2002.

XX PD 12-OCT-2001; 2001WO-US031874.

XX PF 12-OCT-2000; 2000US-0240117P.

XX PR (FERR) FERRING BV.

XX QI S, Akinsanya KO, Riviere PJ, Junien J;

XX PA WPI; 2002-444178/47.

XX DR P-PSDB; ABG61591.

XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding

PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and

PT viral infections, cancers, allergies, neurological disorders, or pain.

XX

PS Claim 1; Page 53-54; 113pp; English.

XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC disease, acute heart failure, hypotension, hypertension, bulimia, Parkinson's
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPPP proteins

XX SQ Sequence 2671 BP; 805 A; 524 C; 594 G; 748 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 2671
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-825-632-1 (1-882) x ABK83322 (1-2671)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

Db 8 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGAAACTGGCGACTGT 67

Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40

Db 68 GAGGAGATATTGAATTCACAGATCGGCTAAATGGAGCTTTTATGTGAGCGGTAT 127

Qy 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

Db 128 TCCTGGAGTCAGCTTAAAAAGCTGCTGCCATACCAAGAAATATCATGCGTACATGATG 187

Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

Db 188 GCTTAAGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 247

Qy 81 AspArgIleTyrTyrIleAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100

Db 248 GACAGAAATCTATTACCTTGCCATGCTGGTGAAGAACAGAGAAATACACTGTTTATCT 307

Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120

Db 308 GAATTTCCCAAACTATCAATAGACAGCAGCTTAAATGCTCTCTTGGAGCCCTCTTTTG 367

Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

Db 368 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAA 427

Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

Db 428 GAAAGAAAAACGCAATTGGAAACAGTCGGAATGTCTTTACGATTTATCCCAAGGAAGTGA 487

Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

Db 488 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCATCGTAAAGATGGAGGCCACACAG 547

Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

Db 548 ATTTACGCAACACCTTTAAGGCCAATCTAGTGGAAACTAGTGTGCCAACATACATCG 607

Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220

Db 608 GATCAAAATATTATGCTCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT 667

QY	221	TpIleSerAenIleValThrArgGluGluArgLeuThyTyxValHisAsnGluLeu	240	Db	1748	GIIGTCCCTTTCAAGCTATCAAGTCTCGAAGATGACCCCACTTGCAAAACAAGGAATTT	1807
Db	668	TGGATATCTAAACATCGTAACAGAGAAGAAGAGACTCACATTATGTGCACATGAGCTA	727	QY	601	TpAlaIleThrIleLeuAspSerAlaGlyProLeuProAspTyxThrProProGluIlePhe	620
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaIleThrPheValLeuGlnGlu	260	Db	1808	TGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTCTCGACTATACTCTCCAGAAATTTTC	1867
Db	728	GCCAAACATGGGAAGAATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGA	787	QY	621	SerPheGluSerThrThrGlyPheThrLeuTyxGlyMetLeuTyxIleHisAspLeu	640
QY	261	PheAspArgTyxSerGlyTyxTTPCyepProLyAlaGluThrThrProSerGlyGly	280	Db	1868	TCCTTTTGAAGTACTACTGGATTACATTTGTATGGAGATGCTCTACAAGCCTCAATGATCA	1927
Db	788	TTTGTAGATATCTCGGCTATTGGTGGTGTCCAAAAGCTGAAACAACCTCCCAAGTGGTGGT	847	QY	641	GlnProGlyLyLeuTyxProThrValLeuPheIleTyxGlyGlyProGlnValGlnLeu	660
QY	281	LysIleLeuArgIleLeuTyxGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Db	1928	CAGCCTGGAAGAAATATCTCTACTGTCTGTTCATATATGGTGGTCTCTCAGGTGCAGTTG	1987
Db	848	AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTCAGGTGGAATTTATTCATGTT	907	QY	661	ValAsnAsnArgPheLeuGlyValIleTyxPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyxProLysThrGlyThr	320	Db	1988	GTGAATAATTCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCTAGCTCTCTAGGT	2047
Db	908	ACATCCCTCTATGTTGGAACAAGGAGGCGAGATTCATTCCGTTATCTTAAACAAGGTACA	967	QY	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
QY	321	AlaAsnProLysValThrPheLeuMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	2048	TATGTGGTTGTAGTGATAGACAACACAGGGATCTCTGTCCCGAGGGCTTAAATTTGAAGGC	2107
Db	968	GCAATCTTAAGTCACCTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAAGGATC	1027	QY	701	AlaPheLeuTyxTyxLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleIleuPheGluGlyValGlu	360	Db	2108	GCCTTTAATAATAAATGGGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATAT	2167
Db	1028	ATAGATGTCATAGATAGAGAACTAAATTCACCTTTTGAGATCTTATTGAGGAGTTGAA	1087	QY	721	LeuAlaSerArgTyxAspPheIleAspLeuAspArgValGlyIleHisGlyTyxTpsSerTyx	740
QY	361	TyrIleAlaArgAlaGlyTTPThrProGluGlyLysTyxAlaIleThrSerIleLeuLeuAsp	380	Db	2168	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT	2227
Db	1088	TATATTTGCCAGACTGGATGACTCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1147	QY	741	GlyGlyTyxLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Db	2228	GGAGGATACCTCTCCCTGATGGCAATATGAGAGGTTCAGATATCTTCAGGGTTGCTATT	2287
Db	1148	CGCTCCCAAGCTCGCTCGAGATAGTGTTCATCTCACCTGAAATATTTATTCACGATAGAA	1207	QY	761	AlaGlyAlaProValThrLeuTTPIlePheTyxAspThrGlyTyxThrGluArgTyxMet	780
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Db	2288	GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATG	2347
Db	1208	GATGATGTTATGGAAAGCAGAGACTCAITTGAGTCAGTGCCTGATTCGTGAGCGCACATA	1267	QY	781	GlyHisProAspGlnAsnGluGlnGlyTyxTyxLeuGlySerValAlaMetGlnAlaGlu	800
QY	421	IleIleTyxGluGluThrThrAspIleTTPIleAsnIleHisAspIlePheHisValPhe	440	Db	2348	GGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2407
Db	1268	ATTATCTATGAAGAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT	1327	QY	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	Db	2408	AAAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTACATGGTTTCTCTGGATGAGAATGC	2467
Db	1328	CCCCAAAGTCACGAAGAGGAATTTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC	1387	QY	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyxAsp	840
QY	461	ArgHisLeuTyxLysIleThrSerIleLeuLysGluSerLysTyxLysArgSerSerGly	480	Db	2468	CATTTTGCACATACCAGTATATTTACTGAGTTTTTTAGTGAGGGCTGGAAGCCATATGAT	2527
Db	1388	CGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1447	QY	841	LeuGlnIleTyxProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyx	860
QY	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	Db	2528	TTACAGATCTATCTCCAGAGACACACAGCATTAAGAGTTCTCTGAATCGGGAGAACATTAT	2587
Db	1448	GGGCTGCCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAGAGAGATAGCAATTTACCAGT	1507	QY	861	GluLeuHisLeuLeuHisTyxLeuGlnGluAsnLeuLysSerArgIleAlaAlaLeuLys	880
QY	501	GlyGluTTPGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	520	Db	2588	GAACTGCATCTTTTGCACTACTCTTCAGAAAAACCTTGATCAGGTATTTGCTGCTCTAAA	2647
Db	1508	GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTTAATATCAAGTTGATGAAGTCAGAAGG	1567	QY	881	ValIle 882	
QY	521	LeuValTyxPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyxValValSer	540	Db	2648	GTGATA 2653	
Db	1568	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTGAAGCATCACCCTGTACGTAGTCAGT	1627	QY			
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyxSerHisSerCysCys	560	Db			
Db	1628	TACGTAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1687	QY			
QY	561	IleSerGlnHisCysAspPheIleSerLysTyxSerAsnGlnLysAsnProHisCys	580	Db			
Db	1688	ATCAGTCAGCAGCTGTGACTTCTTTATAAGTAAGTATAGTAACCCAGAGAATCCACACTGT	1747	QY			
QY	581	ValSerLeuTyxLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	QY			

RESULT 4

ABK12892
ID ABK12892 standard; cDNA; 3106 BP.

XX ABK12892;

XX 09-APR-2002 (first entry)

XX Human protease PRTS-9 cDNA sequence.

XX

KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW cell proliferative disorder; developmental disorder; epilepsy;
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
 KW reproductive disorder; endometriosis; ss.
 XX Homo sapiens.

XX Key Location/Qualifiers
 PH CDS 203..2851
 FT /tag= a
 FT /product= "Human protease PRTS-9"

XX W0200198468-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US019178.

XX 16-JUN-2000; 2000US-0212336P.

XX 22-JUN-2000; 2000US-0213955P.

XX 29-JUN-2000; 2000US-021396P.

XX 07-JUL-2000; 2000US-0216821P.

XX 14-JUL-2000; 2000US-0218946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
 PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
 PI Wallia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
 PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
 PI Kallick DA;

XX WPI; 2002-090437/12.

XX P-PSDB; AAU74749.

XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
 PT (e.g. cancer) disorders.

XX Claim 5; Page 166-167; 177pp; English.

XX The present invention relates to twenty one new human proteases, referred
 CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
 CC invention are useful in the diagnosis, treatment and prevention of
 CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
 CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
 CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
 CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
 CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
 CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
 CC disease and reproductive e.g. infertility and endometriosis disorders.
 CC Numerous other examples of each disorder are given in the specification.
 CC The present nucleic acid sequence encodes the human protease PRTS-9
 CC protein of the invention

XX SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 0 Length: 3106
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-825-632-1 (1-882) x ABK12892 (1-3106)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

Db 203 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGGATATTGGAACTCGCGACTGT 262

Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 263 GAGGAGAAATATTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGGAGCGTAT 322
 Qy 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 323 TCTGTGAGTCAGCTTAAANAAGCTCTTCCCATACCAAGAAATATCATGCTACATGATG 382
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 383 GCTAAGGCACACATGATTTTCTGTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 442
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 443 GACAGAAATCTATTACCTTGCCTATGCTGTGAGAACAGAGAAATATCATGCTACATGATG 502
 Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrTyrLysProLeuLeu 120
 Db 503 GAAATTTCCCAAAATATCAATAGACAGCAGCTTAAATGCTCTCTTGGAAAGCCTCTTTTG 562
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 Db 563 GATCTTTTTCAGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAAAGA 622
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 623 GAAAGAAAACCGCAATTGGAACAGTCGGAATTCCTTACGATATATCATCCCAAGGAAGTGA 682
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 683 ACATTTCTCTTTTCAAGCCGCTAGTGAATTTATCACGTAAGAGATGAGGCGCCACAAGGA 742
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 743 TTTTACGCAACAACTTTTAAAGGCCCAATCTAGTGAACACTAGTTCTCCACATACGATG 802
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 803 GATCCAAATATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 862
 Qy 221 TrpIleSerAsnIleValThrArgGluArgLeuThrTyrValHisAsnGluLeu 240
 Db 863 TGGATATCTAAACATCGTAACACAGAGAAGAAAGAGACTCACTTATGTGCAATAGACTA 922
 Qy 241 AlaAsnMetGluGluAsnAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 923 GCCAAATGGAAGAGATGCCAGATCAGTCGAGTCGCTACTTTTGTCTTCCAAAGAAAGAA 982
 Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 Db 983 TTTGTAGATATTCTGGCTATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1042
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 Db 1043 AAAATTTCTTAGAATTCATATGAGAAATGATGAATCTGAGGTGGAAATTTATTTCATGTT 1102
 Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 1103 ACATCCCTTATGTTGAAACAAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 1162
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 1163 GCAAAATCTTAAGTCACTTTTAAGATGTGAGAAATATGATGCTGCTGAAGGAAGATC 1222
 Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1223 ATAGATGTCATAGATAAGGAACATAATTTCAACCTTTTGAGATTCATTATTGAAGGAGTTGAA 1282
 Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1283 TATATTGCCAGAGCTGGATGGACTCTCTCGGGGAAATATGCTGTGGTCCATCTACTAGAT 1342

QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Db	2423	GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATT	2482
Db	1343	CGCTCCAGACTCGCCTACAGATAGTGTGGATCTCACCTGGAATATTATATCCAGTAGAA	1402	QY	761	AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Db	2483	GTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATG	2542
Db	1403	GATGATGTATTGGAAAGCGAGAGACTCATTTGAGTTCAGTGCCTGATTTCTGTGACGCCACTA	1462	QY	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrIleuGlySerValAlaMetGlnAlaGlu	800
QY	421	IleIleTyrGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe	440	Db	2543	GGTCACCTCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCATGCAAGCAGAA	2602
Db	1463	ATTATCTATGAAGAAACAACAGACATCTGGATAATATCATGACATCTTTTCATGTTTTT	1522	QY	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysIysThrGlyPhe	460	Db	2603	AAGTTCCCTCTGAAACCAATCGTTTACTCTCTTACATCGTTTCTCGGATGGAATGTC	2662
Db	1523	CCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTTTGCCTCTGAAATGCAAAACAGGTTTC	1582	QY	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
QY	461	ArgHisLeuTyrValIleThrSerIleLeuValGluSerIysTyrIysArgSerGly	480	Db	2663	CATTTTGCACATACCAGTATATTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGAT	2722
Db	1583	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1642	QY	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
QY	481	GlyLeuProAlaProSerAspPheIysCysProIleIysGluIleAlaIleThrSer	500	Db	2723	TTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATTAT	2782
Db	1643	GGGCTGCCTCTCCCAAGGATTTCAAGTGTCCTATCAAGAGGAGATAGCAATTACCACT	1702	QY	861	GluLeuHisLeuLeuHisTyrLeuGlnLeuAsnLeuGlySerArgIleAlaAlaLeuLys	880
QY	501	GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	520	Db	2783	GAACTGCATCTTTTGCACTACCTTCAAGAAAACCTTGCATCACGATTGCTGCTCTAAAA	2842
Db	1703	GGTGAATGGGAAGTTCTTGGCGGCATGGATCTTAATATCCAGTTGATGAAGTCAGAAG	1762	QY	881	ValIle	882
QY	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540	Db	2843	GTGATA	2848
Db	1763	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTGAAGCATCACCTGTACGTAGTCAGT	1822	RESULT 5			
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	ID	AAC85694	standard; cDNA; 3120 BP.	
Db	1823	TACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1882	XX	AAC85694;		
QY	561	IleSerGlnHisCysAspPheIleSerIysTyrSerAsnGlnLysAsnProHisCys	580	DT	29-JUN-2001	(first entry)	
Db	1883	ATCAGTCAGCATGTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAAATCCACACTGT	1942	DE	Nucleotide sequence of human DPP8.		
QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysIysThrLysGluPhe	600	XX	Human; dipeptidyl aminopeptidase; DPP8; prollyl oligopeptidase;		
Db	1943	GTGTCCTTTTCAACCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGGAATTT	2002	KW	dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;		
QY	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	KW	growth hormone deficiency; glucose level; mucosal regeneration;		
Db	2003	TGGGCCACCATTTTGGATTCAGCAGGTCCTCTCTGACTATACTCTCCAGAAATTTTC	2062	KW	non-insulin dependent diabetes mellitus; glucose intolerance;		
QY	621	SerPheGluSerThrThrGlyPheThrIleuTyrGlyMetLeuTyrLysProHisAspLeu	640	XX	Homo sapiens.		
Db	2063	TCCTTTGAAAGTACTACTGGATTTTACATTTGATGGAGTCTCTCAAGGCTCATGATCTA	2122	XX	Key Location/Qualifiers		
QY	641	GlnProGlyLysGlyTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660	FT	214..2862		
Db	2123	CAGCCTGGAAAGAAATATCCCTACTGTGCTGTTCATATATGGTGGTCTCAGGTGCAGTTG	2182	FT	/*tag= a		
QY	661	ValAsnAsnArgPheLysGlyValIysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680	FT	/product= "Human DPP8"		
Db	2183	GTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCAGT	2242	XX	W0200119866-A1.		
QY	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700	XX	22-MAR-2001.		
Db	2243	TATGTGTTGTAGTGATAGACAACAGGGGATCCTGTCCAGGAGGCTTAAATTTGAAGGC	2302	XX	11-SEP-2000; 2000WO-AU001085.		
QY	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720	XX	10-SEP-1999; 99AU-00002762.		
Db	2303	GCCTTTAAATATAAATGGGTCAAAATAGAAATTCACGATCAGGTGGAAGACTCCCAATAT	2362	XX	18-FEB-2000; 2000AU-00005709.		
QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheTyr	740	PA	(UNSY) UNIV SYDNEY.		
Db	2363	CTAGCTTCTCGATATGATTTTCATTGCTAGATCGTGGGCATCCACGGCTGGTCTCAT	2422	XX	Abbott CA, Gorell MD;		
QY	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760	XX	WPI; 2001-281520/29.		
				XX	P-PSDB; AAB47187.		
				DR	New human dipeptidyl aminopeptidase (DPP8) useful for cleaving		
				PT	substrates, identifying inhibitors of DPP8 catalytic activity which have		
				PT	therapeutic uses, and for detecting activated T cells.		

XX

Claim 16; Fig 2; 78pp; English.

This sequence encodes human dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase, because it is capable of hydrolyzing the peptide bond C-terminal to Pro in each of these compounds. DPP8 is homologous with human DPP4V. DPP8 is useful for cleaving a substrate, and for detecting an activated T cell which involves measuring the level of DPP8 gene expression in a T cell. The level of DPP8 expression is detected by detecting the amount of DPP8 RNA in the cell. It is also useful for identifying a molecule capable of inhibiting the cleavage of the substrate by DPP8. Molecules identified as inhibiting DPP8 catalytic activity may be useful for treating diarrhoea, inhibiting DPP8 catalytic activity may be useful for treating diarrhoea, growth hormone deficiency, lowering glucose levels in non-insulin dependent diabetes mellitus and other disorders involving glucose intolerance, enhancing mucosal regeneration and as immunosuppressants

XX Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;
SQ

Alignment Scores:

Pred. No.: 0 Length: 3120
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-1 (1-882) x AAC85694 (1-3120)

QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGGAACCTGGGACTGT 273
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 274 GAGGAGAAATTGGAATACAGATCGCGCTTAAATGGAGCCTTTTATGTGTGAGCGGTAT 333
QY 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCCTGGAGTCAGCTTAAAGAGCTGTTGCCATACAGAAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTC 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAAATCTATTACCTTGCCATGTGTGTGAGAACAGAGAAATACACTGTTTATCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTCACAAACATATCAATAGACAGCAGCTTAATGCTCTCTTGGAGCCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 574 GATCTTTTCAGCAACACTGGACTATGAAATGATTTCTCGAGAGAGAACTATTAAAG 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAAACGCAATGGAAACAGTCGGAATTCCTTACGATTTATCACCAGGAAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTCTTCAAGCCGGTAGTGAATTTATCAGGTAAAGATGGAGGGCCACAAAG 753
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTACGCCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCACATACCG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 814 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATACAGCAACG 873

QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
DB 874 TGATATCTAAACATCGTAACAGAGAGAAAGAGAGACTCACTTATGTGCACATGAGCTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
DB 934 GCCAACATCGAAGAGATGCCAGATCAGCTGGAGTCGTACTCTTTGTTCTCCAAAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATCTGGCTATTGGTGTGTCACAAAGCTGAACAACTCCCAAGTGGTGT 1053
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
DB 1054 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCAT 1113
QY 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTTATGTTGGAACAAGAGGCGCAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 1174 GCAATCTCTAAAGTCACITTTTAAGATGTGAGAAATATGATGCTGAAGGAAGGATC 1233
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGTGAGATTCATTATTTGAAGGAT 1293
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTGTGCTCACTACTAGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
DB 1414 GATGATGTTATGGAAAGGCAGAGACTCATTTGAGTCAGTCGCTGATTTCTGTGACGCCACTA 1473
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1474 ATTATCTATGAAGAACCAACAGACATCTGGATAATATCCATGACATCTTTTCATGTTT 1533
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1534 CCCAAAGTCACGAGAGGAAATTCAGTTTATTTTGTCTCTGATGCATAAACAGGTTTC 1593
QY 461 ArgHisLeuTyrIleSerIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
DB 1594 CGTCATTTATACAAATTAACATCTATTTTAAAGGAAGCAAAATATAAACGATCCAGTGGT 1653
QY 481 GlyLeuProAlaProSerAspPheLysCysProLysLeuGluGluIleAlaIleThrSer 500
DB 1654 GGGCTCGCTGCTCCAGATGATTTCAAGTGTCTATCAAAAGAGGAGATAGCAATTAACCA 1713
QY 501 GlyValTrpGluValLeuGluValArgHisGlySerAsnIleGlnValAspGluValArgArg 520
DB 1714 GGTGAATGGGAAGTCTCTTGCCGGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1774 CTGATATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TAGCTAAATCTCGAGAGGTGACAAGGCTGACACCGTGGCTACTCACATTTCTTGTGTC 1893
QY 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACCAAGAGAAATCCACACTGT 1953

101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProIleuLeu 120
114 GAAATTCACAAATATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTG 573
121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140
134 GATCTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAAGAAGAACTATTAGA 633
141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
154 GAAAGAAACGCAATTTGGAACAGTGGAAATGCTTCTTACCAATATCACCAGAGAGTGA 693
161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
174 ACAATTTCTGTTTCAAGCGGTAGTGAATTTATCACGTAAGAATGAGAGGCGGCACAGA 753
181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
194 TTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATG 813
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
214 GATCCAAATTAATGCGCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATAT 873
221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
234 TGGATATCTCAATCGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
254 GCCAAACATGGAG 993
261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
274 TTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGT 1053
281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
294 AAAATCTTAGAATCTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1113
301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
314 ACATCCCTCTATTTGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1173
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluArgIle 340
334 GCAATCTCTAAAGTCACTTTTAAGATGTGAGAAATATGATGCTGAAGAGAGATC 1233
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
354 ATAGATGTCTAGATAAG 1293
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
394 TATATTCAGAGAGTGGATGAGTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1353
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
394 CGCTCCCAAGACTCGCTTACAGATGTGTGATCTCACCTGAATTTATTTATCCAGAGAG 1413
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
414 GATGATGTTATGGAAG 1473
421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
454 ATTATCTATGAAGAAACACAGACATCTGGATAATATCCATGACATCTTTTCATGTTTT 1533
441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
474 CCCCAGAGTCAAG 1593

461 ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
474 CGTCATTTTATCAAAATATACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT 1653
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
494 GGGCTCGCTGCTCCAAAGTGTCTTCAAGTGTCTTCAAGAGAGAGATAGCAATTTACCAGT 1713
501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
514 GGTGAATGGGAAGTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValLysSer 540
534 CTGGTATATTTTGAAGGACCAAGAGCTCCCTTTTAGAGCATCACCTGTAGTAGTCAGT 1833
541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
554 TAGTAAATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTC 1893
561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
594 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACCAAGAGAAATCCACACTGT 1953
581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
594 GTGTCCCTTTTACAAGCTATCAAGTCTGAAGATGACCACTGTCNAAAACAAAGAGAAATTT 2013
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
614 TGGGCCACCATTTTGGATTCAGCAGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
654 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTA 2133
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
674 CAGCCTGGAAAGAAATATCTACTGTGCTGTTTATATATGGTGTCTCTCAGGTGTCAGTTG 2193
661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
694 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCCTTGAATATCCCTAGCCTCTCTAGGT 2253
681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
714 TATGTGTTGTAGTAGACACAACAGGGGATCTGTCAACCGGCGCTTAAATTTGAAGGC 2313
701 AlaPheLysTyrIleMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyr 720
734 GCCTTTAAATATAAAATGGGTCAAAATAGAAAATTCAGCATCAGGTGGAGGAGCTCCAATAT 2373
721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
754 CTAGCTTCTCGATATGATTTCAATGACTGTAGTGTGGGATCCACCGCTGTCTCTAT 2433
741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
774 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGTTGCTATT 2493
761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
794 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
814 GGTCAACCTGACCAAGTGAACAGGCTTACTTACTAGGATCTGTGGGATGTCAGAGAGAA 2613
801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
834 AAGTTCCTCTGGAACCAATCTTTACTTCTTACTTCTTACTTCTTCTTCTTCTTCTTCTTCT 2673
821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

DB 2674 CATTTGCACATACAGATATATTACTGAGTTTTTTAGTGAGCGCTGGAAACCCATATGAT 2733
QY 841 LeuGlnIleTyrProGlnGluA:GHisSerIleArgValProGlnSerGlyGluHisTyr 860
DB 2734 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGACATAT 2793
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluA:SerArgIleAalaLeuLys 880
DB 2794 GAACTGCATCTTTTGCACCTACCTTCAGAAAACCTTTGGATCAGTATTGCTGCTTAAAA 2853
QY 881 ValIle 882
DB 2854 GTGATA 2859
AAH99934
ID AAH99934 standard; cDNA; 3143 BP.
XX
AC AAH99934;
XX
DT 12-APR-2002 (first entry)
DE cDNA encoding 21953 human prolyl oligopeptidase.
KW 21953 prolyl oligopeptidase; human; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 229..2877
XX /*tag= a
XX /product= "21953 prolyl oligopeptidase"
XX /note= "This region is specifically claimed in claim 2"
XX
XX WO200179473-A2.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US040483.
XX
XX 18-APR-2000; 2000US-0197508P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers RA, Williamson M;
XX
XX WPI; 2002-034353/04.
XX P-PSDB; AAG78415.
XX
XX New polypeptides 21953, member of human prolyl oligopeptidase family,
XX useful as diagnostic targets and therapeutic agents for controlling
XX cancer, lymphoma and leukemia.
XX
XX Claim 7; Page 100-102; 121pp; English.
XX
XX This invention relates to an isolated 21953 human prolyl oligopeptidase.
XX Which is cytosolic, antidiabetic, antiarthritic, neuroprotective,
XX antithyroid, dermatological, antipsoriatic, antiasthmatic,
XX ophthalmological, antiinflammatory, nontropic, antiparkinsonian,
XX anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,
XX antithrombotic, anorectic and metabolic in its action. Uses include
XX gene therapy, expression or activity of 21953 protein modulator, it is
XX useful for identifying a compound which binds to it and can be used in
XX preventing, treating or detecting a cellular proliferative or
XX differentiative disorder. The 21953 molecules can act as novel diagnostic
XX targets and therapeutic agents for controlling disorders associated with
XX the aberrant activity or degradation of peptide hormones e.g., disorders

CC associated with cell differentiation and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the cDNA encoding
CC sequence of 21953 human prolyl oligopeptidase
XX
XX Sequence 3143 BP; 943 A; 644 C; 712 G; 844 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 3143
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 6
US-10-825-632-1 (1-882) x AAH99934 (1-3143)
QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 229 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTCGCGACTGT 288
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 289 GAGGAGATATTTGAATACAGATCGCGCTAAATTTGGAGCTTTTATGTGAGCGGTAT 348
QY 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 349 TCTTGAGTCAGCTTAAAGAGCTTTCCTGCGATACCAAGAAATATCATGCTACATGATG 408
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 409 GCTAAGCGCACCACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTC 468
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 469 GACAGAAATCTATTACCTTGCCTATGCTGTGTGAGAACAGAGAAATACACCTGTTTTATCT 528
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 529 GAAATTTCCAAATCTATCAATAGACGACAGCTCTTAATGCTCTCTTGGAAAGCCTCTTTTG 588
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 589 GATCTTTTTCAGCAACACTGCACTATGCAATGTATTCTCGAGAAGAAGAACTATTAAAG 648
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 649 GAAAGAAAACCGATTGGACACAGTCGGAATTCCTTACGATTATCACCAAGGAAGTGA 708
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 709 ACATTTCTGTTTCAAGCCGTAGTGAATTTATCACTGTAAGAATGAGAGGCCACAGGA 768
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleAspMet 200
DB 769 TTTTACGCAACAACTTTTAAAGGCCCAATCTAGTGAACACTAGTTGTCCACATACGATG 828
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 829 GATCCAAAAATTTATGCCCTGCTGATCCAGACTGGATGCTTTTATACATAGACAGATAT 888
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
DB 889 TGGATATCTAAATCTGTAACCAAGAGAAGAAAGAGACTCACTTATGTGCAATAGAGCTA 948
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 949 GCCAACATGGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAGAGAA 1008

Qy 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 1009 TTTGATAGATATTTGGCTATTGGTGTGCCAAAGCTGAACCACTCCAGTGGTGT 1068
Qy 281 LysIleuArgIleuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 1069 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATCATGTT 1128
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1129 ACATCCCTATGTTGGAAACAGAGGAGGAGCAGATTCATTCGTTATCTTAAACACAGGTACA 1188
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1189 GCAATCTCTAAAGTCACTTTTAAAGATGTGAGAAATATGATGATGCTGAAGGAAGGATC 1248
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluValGlu 360
Db 1249 ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGGAGATTTCTATTGGAAGAGTTGAA 1308
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1309 TATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGGTCCATCTCTACTAGAT 1368
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1369 CGCTCCAGACTCGCTCGAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1428
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1429 GATGATGTTATGGAAGGAGAGAGACTCATTTGAGTCAGTCGCTGATTTCTGTGAGCCCACTA 1488
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1489 ATTATCTATGAAGAAACACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1548
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1549 CCCAAAGTCACGAGAGGAATTTGATTTATTTGCTCTGATGCAAAACAGGTTTC 1608
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1609 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACAGATCCAGTGT 1668
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1669 GGGCTGCTCTCCAAAGTATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1728
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValAlaArg 520
Db 1729 GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAATCAGNAGG 1788
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1789 CTGGTATTTTGAAGGCCAACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1848
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1849 TACGTAATCTCGAGAGGTGACAGAGGCTGACTACCGCTGGCTACTCACATTTCTTGCTGC 1908
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1909 ATCAGTCAGCACTGTGACTTCTTTATAGTAGTAGTAGTAGTACCAAGAGATCCACACTGT 1968
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1969 GTGTCCCTTTACAGCTATCAAGTCTGAGATGACCAACTTGCACAAACAAAGGAATTT 2028
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
Db 2029 TGGGCCCACTTTTGGATTTCAGCAGGTCTCTTCTGACTATCTACTCTCCAGAAATTTTC 2088

Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2089 TCTTTTGAAGTACTACTGGATTTTACATGTATGGATGCTCTACAAGCTCTATGATCTA 2148
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2149 CAGCTTGAAGAAATATCTCTACTGCTGCTTATATATATGCTGCTCTCAGGTGAGTTG 2208
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2209 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGT 2268
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2269 TATGTGGTGTAGTAGATAGACCAACAGGGGATCCTGTACCAGGGCTTTAAATTTGAAGC 2328
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2329 GCCTTTAATATATAATGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCATAT 2388
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2389 CTAGCTTCTCGATATGATTTTCACTTAGATCGTGTGGCATCCACGGCTGCTCTAT 2448
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2449 GAGGAGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTTCAGGGTTGCTATT 2508
Qy 761 AlaGlyValProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2509 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGACGTTATATG 2568
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2569 GGTCACTCCCTGAGCAAGTCAACAGGGCTATTACTTAGGATCTGTGGCATCGAAGCAGAA 2628
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2629 AAGTTCCTCTCGAACAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC 2688
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2689 CATTTGCAATACCACTAGTATATTACTGAGTTTTTTAGTGGGGCTGGAAAGCCATATGAT 2748
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2749 TTACAGATCTATCTCTCAGAGAGACACAGATAGAGTTCTGTAATCGGAGAACATTAT 2808
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2809 GAACGTGCATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCACGTATTGCTGCTCTAAA 2868
Qy 881 ValIle 882
Db 2869 GTGATA 2874
RESULT 8
AAH99935
ID AAH99935 standard; cDNA; 2643 BP.
XX
AC AAH99935;
XX
DT
XX 12-APR-2002 (first entry)
DE Coding sequence of 21953 human prollyl oligopeptidase.
XX 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW antidiabetic; antiarthritic; antiaesthetic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease; ss.
XX

CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
XX -ABK83343 encode human DRP proteins
SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 4829
Score: 4680.00 Matches: 882
Percent Similarity: 99.8% Conservative: 0
Best local Similarity: 99.8% Mismatches: 0
Query Match: 99.6% Indels: 2
DB: Gaps: 6

US-10-825-632-1 (1-882) x ABK83327 (1-4829)

QY	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
DB	214	ATGGCAGCAGCAATGGAAACAGAACACAGCTGGGTGTGAGATATTTGAAACTGCCGACTGT	273
QY	21	GluGluAenilleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
DB	274	CAGGAGAAATATTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATATTGAGCGGTAT	333
QY	41	SerTtpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
DB	334	TCCTGGACTCAGCTTAAAGCTGCTTGCAGATACACAGAAATATCATGGCTACATGATG	393
QY	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
DB	394	GCTAAGGCCACCATGATTTTCATGTTTGTCAAGAGGAATGATCCAGATGGACCTCATTTCA	453
QY	81	AspArgLileTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
DB	454	GACGAATCTATTACCTTGCATGCTCTGGTGAGACAGACAGAAATACACTGTTTTATTCT	513
QY	101	GluileProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
DB	514	GAAATTCACAAACATCATATAGACAGACCTTAATGCTCTCTTGGAGCCTCTTTTG	573
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
DB	574	GATCTTTTTCAGGCACACTGGACTATGGAAATTCCTTCCAGAGAGAGACTATTNAGA	633
QY	141	GluArgLysArgLileGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
DB	634	GAAAGAAAACGCAATTTGGAAACAGTCGGAAATTCCTTACCAGATATCACAGGAAGTGA	693
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
DB	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAGGA	753
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAenilleArgMet	200
DB	754	TTTACGCAACAAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG	813
QY	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
DB	814	GATCCCAAAATATGCGCTGCTCATCCAGACTGGATGCTTTATATACATAGCAACGATATT	873
QY	221	TrpIleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeu	240
DB	874	TGATATCTAAACATCGTAACACAGAGAGAGAGAGAGACTCATTTATGTGCAACATGAGCTA	933
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
DB	934	GCCACATATGGAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA	993
QY	261	PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly	280
DB	994	TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT	1053
QY	281	LysIleLeuArgLileLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300

DB	1054	AAAATCTTGAATAATCTATATGAGAGAAATGATGAATCTGAGGTGGAAATATTTCATGTT	1113
QY	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrTyrProLysThrGlyThr	320
DB	1114	ACATCCCTCATGTTGGAAACCAAGAGGCGAGATTATCCGTATATCTTAAACAGGTACA	1173
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
DB	1174	GCAAAATCTTAAAGTCATCTTTTAAAGATGTGAGAAATAATGATGCTGAGGAAGGATC	1233
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
DB	1234	ATAGATGTCAATAGAGAACTAAATCAACCTTTTGAGATTCTATTTTGAAGGAGTTGAA	1293
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
DB	1294	TATATTGCCAGAGCTGGATGCACTCTCTGAGGGAAAAATATGCTTGGTCCATCTCTACTAGAT	1353
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
DB	1354	CGCTCCAGACTCGCTTACAGATAGTGTGTGATCTCACCTGAATATTATTATCCAGTAGAA	1413
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
DB	1414	GATGATGTTATGGAAAGCGCAGAGACTCATTTGACTGCTGCTGATTTCTGTGACGCCACTA	1473
QY	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleIleHisAspIlePheHisValPhe	440
DB	1474	ATTATCTATGAGAAACACACAGCATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533
QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
DB	1534	CCCCAAAGTCACGAAGAGAAATTTGAGTTATTTTTTGGCTCTGATGCAAAACAGGTTTC	1593
QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
DB	1594	CGTCATTTTATCAAAATTTACATCTATTATTAAGAAAGCAAAATATTAACCATCGGTGT	1653
QY	481	GlyLeuProAlaProSerAspPheLysCysProLleLysGluIleAlaIleThrSer	500
DB	1654	GGGCTCGCTCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTCAGT	1713
QY	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
DB	1714	GGTGAATGGAAAGTTCTTTGGCCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAG	1773
QY	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValSer	540
DB	1774	CTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCATCCTGTACGTAGTCAGT	1833
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
DB	1834	TAGTAAATCTTGGAGAGTGCAAGGCTGACTCACCGTGGCTACTCACATTTCTTGTCTGC	1893
QY	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
DB	1894	ATCAGTCAGCATGTGACTTCTTTATAGTAAGTATAGTACCAAGAAATCCACACTGT	1953
QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
DB	1954	GTGTCCCTTTACAGCTATCAAGTCTCTGAGATGACCACTTGCAGAAACAAAGGAATTT	2013
QY	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
DB	2014	TGGGCCACCAATTTTGGATTCAGAGTCTCTTCTGACTATACTCTCCAGAAATTTTC	2073
QY	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
DB	2074	TCTTTTGAAAGTAGTACTACTGGATTACATTGATGGATGCTCTCAACAGCTCATGACTA	2133
QY	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL	660
DB	2134	CAGCCTGGAAAGAAATATCTCTACTGTGCTGTTTCATATGATGGTGGTCTCTCAGGTGCA	2193


```
QY 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuG 680
DB 2194 TGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCTAGCTCTCTAG 2253
QY 680 lYTyValValValValLysPheAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
DB 2254 GTTATGTGTTGTAGTATAGACACACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAG 2313
QY 700 lYAlaPheLysTyrLysMetGlyGlnLleGluLleAspAspGlnValGluGlyLeuGlnT 720
DB 2314 GGCCTTTTAAATATTAATATGGTCAATAGAAATTCAGATCAGGTGGAGGACCTCCAT 2373
QY 720 YrLeuAlaSerArgTyrAspPheLysPheLeuAspArgValGlyLleHisGlyTyrSerT 740
DB 2374 ATCTAGCTTCTCGATATGATTTTCACTTAGATCGTGTGGCATCCACGGCTGTGCTCT 2433
QY 740 YrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspLlePheArgValAlaI 760
DB 2434 ATGGAGGATACCTCTCCCTGATGCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTA 2493
QY 760 lAlaGlyAlaProValThrLeuTrpLlePheTyrAspThrGlyTyrThrGluArgTyrM 780
DB 2494 TTGCTGGGGCCCGTCACTCTGTGGATCTTCTATGATACGATACACGGACGTTATA 2553
QY 780 etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG 800
DB 2554 TGGGTCACTGACCCAGAAATGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAG 2613
QY 800 lYlAsPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
DB 2614 AAAAGTTCCCTCTCAACCAATCGTTTACTGTCTTTACATGTTTCTCGGTGAGAAATG 2673
QY 820 aHisPheAlaHisThrSerLleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA 840
DB 2674 TCCATTTTGCACATACCAAGTATATTACTGAGTTTTTTTAGTGAGGGCTGGAAAGCCATATG 2733
QY 840 sPLeuGlnLleTyrProGlnGluArgHisSerLleArgValProGluSerGlyGluHisT 860
DB 2734 ATTTACAGATCTATCTCTCAGGAGACACACAGATTAAGAGTTCTGTAATCGGAGAACATT 2793
QY 860 YrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgLleAlaLeuL 880
DB 2794 ATGAAGTGCATCTTTTGCACCTTCAAGAAAACCTTGGATCACGTTATTGCTGCTCTAA 2853
QY 880 ysValIle 882
DB 2854 AAGTGATA 2861
RESULT 10
ADL13374
ID ADL13374 standard; cDNA; 2797 BP.
AC AC
XX ADL13374;
XX ADL13374;
DT 06-MAY-2004 (first entry)
XX Human steroid-induced C3A liver cell cDNA #1103.
DE
XX ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX Homo sapiens.
OS
XX US6673549-B1.
PN
XX 06-JAN-2004.
PD
XX 12-OCT-2001; 2001US-00976594.
XX 12-OCT-2000; 2000US-0240409P.
XX (INCY-) INCYTE CORP.
PA
```

```
XX Furness LM, Buchbinder JL;
PI WPI; 2004-068610/07.
XX Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX Claim 1; SEQ ID NO 1103; 141pp; English.
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2797
Score: 4443.50 Matches: 840
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 1
Query Match: 94.5% Indels: 3
DB: 12 Gaps: 1

US-10-825-632-1 (1-882) x ADL13374 (1-2797)
QY 42 TrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMetAla 61
DB 3 TGGAGTCAGCTTAAAAGCTGCTGCCGATACCGAGAAATATCATGCTACATGCGCT 62
QY 62 LysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSerAsp 81
DB 63 AAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGAC 122
QY 82 ArgLleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSerGlu 101
DB 123 AGAATCTATTACCTTGCCATGCTGTGTGAGAACAGAGAAATATACACTGTTTTTATCTGAA 182
QY 102 IleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeuAsp 121
DB 183 ATTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGGCTCTTTTGGAT 242
QY 122 LeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArgGlu 141
DB 243 CTTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAGAGAA 302
QY 142 ArgLysArgLleGlyThrValGlyLleAlaSerTyrAspTyrHisGlnGlySerGlyThr 161
DB 303 AGAANAACGATTTGGACAGTCGGAATGCTCTTACGATTATACCAAGGAAGTGGAAACA 362
QY 162 PheLeuPheGlnAlaGlySerGlyLleTyrHisValLysAspGlyGlyProGlnGlyPhe 181
DB 363 TTTTCTGTTTCAAGCCGCTAGTGAATTTATCATCGTAAAGATGGAGGGCCACACAGGATT 422
QY 182 ThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAsp 201
```

423	Db		ACGCAACAACTTTAAGGCCCAATCTAGTGGAAACATAGTTGTCCCAACATACGAGTGGAT	482
202	QY		ProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIleTrp	221
483	Db		CCAAATTAATGCCCTCTGATCCAGATGGATTGCTTTTATACATAGCAACGATATTGG	542
222	QY		IleSerAsnIleValThrArgGluGluArgLeuThrTyValHisAsnGluLeuAla	241
543	Db		ATATCTAACATCGTACCAGAGAGNAGGAGACTCACTTATGTGCAATAGCTAGCC	602
242	QY		AsnMetGluGluAspIleArgSerAlaGlyValAlaThrPheValLeuGlnGluPhe	261
603	Db		AACATGAGAGAGATGCCAGATCAGCTGGAGTCGTACTCTTTGTCTCCAGAGAATTT	662
262	QY		AspArgTySerGlyTyTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLys	281
663	Db		GATAGATATTCTCGCTATTGGTGGTGTCCAAAGCTGAACAACTCCCGTGGTGTAA	722
282	QY		IleLeuArgIleLeuTyrgluGluAsnAspGluSerGluValGluIleHisValThr	301
723	Db		ATTCTTAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTAATTCATGTACA	782
302	QY		SerProMetLeuGluThrArg-ArgAlaAspSerPheArgTyProLysThrGlyThrAl	321
783	Db		TCCCTCATGTGTGAAACAAGCGCAGGCAGATTCAATCCGTATTATCTTAAACAGGTACAGC	842
321	QY		aAsnProLysValThrPhelysMetSerGluIleMetIleAspAlaGluGlyArgIleIl	341
843	Db		AAATCTCTAAAGTCACATTTTAAAGATGTCAGAAATAATGATTGATGTGCTGAAGGAAGGATCAT	902
341	QY		eAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTy	361
903	Db		AGATGTCATAGATAAGGAACATAATCCAACCTTTTGAGATTCTATTTTGAAGGAGTTGAATA	962
361	QY		rIleAlaArgAlaGlyTrpThrProGluGlyLys---TyrAlaTrpSerIleLeuLeuA	380
963	Db		TAITGCCAGAGCTGGATGGACTCTCTGAGGGAAAAATAGTGCTGCTGGTCCATCTACTAG	1022
380	QY		spArgSerGlnThrArgLeuGlnIleValIleLeuSerProGluLeuPheIleProValG	400
1023	Db		ATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTAATTTATCCAGTAG	1082
400	QY		luAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProL	420
1083	Db		AAGATGATGTTATGGAAGGACAGAGACTCATTAGTCAGTGGCTGATCTCTGTGACGCCAC	1142
420	QY		euleIleTyrgluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValP	440
1143	Db		TAATTAATCTATGAAGAAACAACAGACATCTGGATAAATAATCCATGACATCTTTCATGTTT	1202
440	QY		heProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyP	460
1203	Db		TTCCCCAAAGTCAGAGAGGAAATTAGTATTATTTTCCCTCTGAAATGCAAAACAGGTT	1262
460	QY		heArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyryAspArgSerSerg	480
1263	Db		TCCGTCATATTATACAAATTAACATCTATTATTAAGAAAGAGCAAAATATAAACGATCCAGTG	1322
480	QY		lyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrS	500
1323	Db		GTGGGCTGCTGCTCCAAAGTATTCAAGTGTCTTATCAAGAGGAGATAGCAATATACCA	1382
500	QY		erGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgA	520
1383	Db		GTGGTGAATGGGAAGTCTTGGCGCGCATGATCTAAATATCCAAGTTGATGAAGTCAGAA	1442
520	QY		rgLeuValTyPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyValValS	540
1443	Db		GGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCA	1502
540	QY		erTyValAsnProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysAc	560

RESULT 11
ACA92421
ID ACA9

XX
AC ACA92421;
XX
DT 15-JUL-2003 (first entry)
XX
DE DNA encoding human PMMM-6.
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cycosatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thyromimetic; gene; ds.
OS
XX Homo sapiens.
XX
PN WO2003031939-A2.
XX
XX 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
XX 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BW, Lee EA, Griffin JA, Li JX;
PI Sprague JW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
XX WPI; 2003-430274/40.
DR P-PSDB; ABU92026.
XX
XX New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
XX Claim 5; Page 285-286; 31pp; English.
XX
XX The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
XX ACA92416-ACA92455 encode the human PMMM polypeptides of the invention
XX
SQ Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2952
Score: 4397.50 Matches: 831

Percent Similarity:	94.2%	Conservative:	0
Best Local Similarity:	94.2%	Mismatches:	51
Query Match:	93.6%	Indels:	1
DB:	10	Gaps:	1
US-10-825-632-1 (1-882) x ACA92421 (1-2952)			
Qy	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	204	ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGAAACCTGGGAGCTGT	263
Qy	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Db	264	GAGAGAAATATTGAATCACAGATCGGCTAAATTTGGAGCCCTTTTATGTGAGCGGTAT	323
Qy	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Db	324	TCCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAGAAATATCATGGCTACATGATG	383
Qy	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Db	384	GCTAAGGCACCATGATTTTCATGTTGTGAAGAGGAATGATCAGATGGACCTCATTTCA	443
Qy	81	AspArgIleTyrTyrIleAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Db	444	GACAGAAATCTATTACCTTGCCATGTCTGGTGAAGACAGAGAAATACACTGTTTATCT	503
Qy	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Db	504	GAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCCTCTTTG	563
Qy	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg	140
Db	564	GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAGA	623
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Db	624	GAAGAAGAAACGCAATTTGGAAACAGTCGGAATGCTCTTACGATTTATCACCAGGAAGTGA	683
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Db	684	ACATTTCTGTTTCAGCGCGTAGTGAATTTATCAGGTAAAGATGGAGGCGCCACAGGA	743
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Db	744	TTTACGCAACACCTTTAAGGCCAATCTAGTGGAAACTAGTGTGCCCAACATACGCGATG	803
Qy	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Db	804	GATCCAAAATTTATGCTGCTGATCCAGCTGGATGCTTTTATACATAGCAACGATATT	863
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu	240
Db	864	TGATATCTAAACATCTGTAACCAAGAGAAGAGAGACTCATTTATGTCACACATAGCTA	923
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
Db	924	GCCAAATATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA	983
Qy	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Db	984	TTTGATAGATATTCTGGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1043
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
Db	1044	ANAATCTTAGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1103
Qy	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Db	1104	ACATCCCTATGTTGGAACAGAGGCGGCGAGATTCATTCCTCCGTTATCTCTAAACAGGTACA	1163
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340

```
Db 1164 CCAAACTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATC 1223
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGluValGlu 360
Db 1224 ATAGATGTCATAGATAAGGAACCTAATATCAACCTTTTGAGATCTATTTTGAAGAGTTGAA 1283
Qy 361 TyrIleAlaArgAlaGlyThrProGluGlyLysThrAlaThrSerIleLeuLeuAsp 380
Db 1284 TATATTGCCAGAGCTGATGGACTCCCTGAGGGGAAATATGCTTGGTGCATCTCTAGAT 1343
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1344 CGCTCCACAGACTCGCTACAGATAGTGTGATCTCACTCGAAATATTTATCCCAAGTAGAA 1403
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1404 GATGATGTTATGGAAGGCGAGACTCAATTGAGTCAGTGCCTGATTCGTGACGCCACTA 1463
Qy 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
Db 1464 ATTATCTATGAAGAACACACAGACATCTGATATAATATCATGACATCTTTTCATGTTTT 1523
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1524 CCCCAAGTCACGAAGAGGAAATTGAGTTATTTTTCCTCTGAATGCCAAACACAGTTTC 1583
Qy 461 ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysLysThrLysArgSerSerGly 480
Db 1584 CGTCATTTATACAAATTTTACATCTATTTTAAAGAAAGCAATATATACGATCCAGTGGT 1643
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1644 GGGCTGCTGCTCCAAAGTGAATTCATAGTGTCCTATCAAGAGAGATAGCAATACCAAGT 1703
Qy 501 GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1704 GGTGAATGGGAAGTCTCTTGCCCGCATGGATCTAAATATCCAAAGTTGATGAAGTCAGAAG 1763
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
Db 1764 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAAAGCATCACCTGTACGTACGT 1823
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1824 TACGTAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTCCTTGCTGC 1883
Qy 561 IleSerGlnHisCysAspPheIleSerLysThrSerAsnGlnLysAsnProHisCys 580
Db 1884 ATCAGTCAGCAGCTGTGACTCTTTTATAAGTAAAGTATAGTAACCCAGAGAATCCACACTGT 1943
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1944 GTGTCCTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAGGAATTT 2003
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2004 TGGGCCACATTTTGGATTACGAGGTCCTCTTCCTGACTATATCTCTCCAGAAATTTTC 2063
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2064 TCTTTTGAAGACTACTGGAATTTTACATTTGATGGGATGCTCTACAAAGCCTCATGATCTA 2123
Qy 641 GlnProGlyLysLysThrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2124 CAGCCTGGAAGAAATATCTCTACTGCTGCTGCTATATATGTTGGTGGTCTCAGGTGAGTTG 2183
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2184 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCGGCTTGAATACCCCTAGCCTCTCTAGT 2243
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2244 TATGTGGTGTAGTGATAGACAAACAGGGGATCTCTGTCCAGGGGCTTAAATTTGAAGGC 2303
```

```
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2304 GCCITTAATATAAATG----- 2321
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2321 ----- 2321
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2322 -----GTTGCTATT 2330
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2331 GCTGGGGCCCGAGTCACCTCTGTGATCTTCTATGATACAGATACACGGAACGTTATATG 2390
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2391 GGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCACAGCAA 2450
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2451 AAGTTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAATGC 2510
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2511 CATTTTGCACATACCAAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGAT 2570
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2571 TTACAGATCTATCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCTGGAGAACATTAT 2630
Qy 861 GluLeuHisLeuLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2631 GAACGTGCAATCTTTGCACTACTTCAAGAAAACCTTGGATCAGCATATTCGTGCTCTAAA 2690
Qy 881 ValIle 882
Db 2691 GTGATA 2696
```

RESULT 12
ABK83332
ID ABK83332 standard; cDNA; 4685 BP.
XX AC ABK83332;
XX DT 12-AUG-2002 (first entry)
XX DE cDNA encoding human DPRP-1 splice variant #8.
XX KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
XX KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
XX KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
XX KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
XX KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
XX KW dyskinesia; reproductive disorder; inflammatory disorder;
XX KW metabolic disorder; gene; ss.
XX OS Homo sapiens.
XX PN WO200231134-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.
XX PR 12-OCT-2000; 2000US-0240117P.
XX PA (PERR) FERRING BV.
XX QI S, Akinsanya KO, Riviere PJ, Junien J;

DR WPI; 2002-444178/47.
 XX P-PSDB; ABG61601.
 PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
 PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
 PT viral infections, cancers, allergies, neurological disorders, or pain.
 XX
 PS Disclosure; Page 75-76; 113pp; English.
 XX
 CC The present invention relates to the isolation of novel human serine
 CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
 CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
 CC acids encoding them are useful for treating infections such as fungal,
 CC bacterial, protozoan and viral infections, particularly infections caused
 CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
 CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
 CC disease, acute heart failure, hypotension, hypertension, urinary
 CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
 CC allergies, cancers, migraine, vomiting, psychotic and neurological
 CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
 CC These may also be used in discovering therapeutic agents for the
 CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
 CC -ABK83343 encode human DPRP proteins
 XX
 SQ Sequence 4685 BP; 1430 A; 853 C; 991 G; 1411 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 4685
 Score: 4395.50 Matches: 834
 Percent Similarity: 94.6% Conservative: 0
 Best Local Similarity: 94.6% Mismatches: 1
 Query Match: 93.3% Indels: 48
 DB: 6 Gaps: 1

US-10-825-632-1 (1-882) x ABK83332 (1-4685)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
 Db 214 ATGCAGCAGCAATGGAACAGACAGCTGGGTGGATATTTGAACTGGGACTGT 273
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 274 GAGGAGAAATATGAAATCACAGGATCGGCTAAATGGAGCCCTTTTATGTGGCGGTAT 333
 Qy 41 SerTrpSerGlnLeuLysIleLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGGATACCAAGAAATATCATGCTACATGATG 393
 Qy 61 AlAlaAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATTC 453
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 454 GACAGAATCTATTACCTTGCATGCTGCTGGTGAAGACAGAGAAATACACATGTTTATCT 513
 Qy 101 GluLeuProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 514 GAAATTCCTCCAAATATCAATAGACAGAGCTTTAAATGCTCTCTTGAAGCCCTCTTTT 573
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 Db 574 GATCTTTTTCAGCAACACATGGACTATGGAATGATTTCTCGAAGAGAAAGAACTATTAGA 633
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 634 GAAAGAAACGCGATGGACAGTCGGAATGCTCTTACGATTATCACCAAGGAAGTGA 693
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 694 ACATTTCTGTTCAGCCGGTAGTGAATTTATCATCGTAAAGATGGAGGGCCACAGGA 753
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

Db 754 TTTACGCAACAACTTTTAAAGCCCACTAGTGGAACTAGTTGTCCACATACCGATG 813
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 814 GATCCAAATTTATGCTGCTGATCCAGACTGGATGCTTTTATATACATACCAACGATATT 873
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 Db 874 TGGATATCTTAACATCGTAACAGAGAGAAAGAGAGACTCACTTATGTGCAATAGACTA 933
 Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 934 GCCAATGGAAGAGATGCCAGATCAGCTGGATCGCTACCTTTGTTCTCCAGAGAGAA 993
 Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 Db 994 TTTGATAGATATTCGCTATTGGTGGTGTCCAAAGCTGAACCAACTCCCAAGTGGTGT 1053
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 Db 1054 AAAATTTCTTAGAATTTCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
 Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 1114 ACATCCCTTATGTTGGAACAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 1174 GCAATCTCTTAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAAGGAGGATC 1233
 Qy 341 IleAspValIleAspLysGluLeuGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1234 ATAGATGTCATAGATAAGGAACATAATTCACCTTTGAGATTCCTATTGAGGAGTTGAA 1293
 Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1294 TATATTGCCAGAGCTGGATGGACTCTGAGGAGAAATATGCTTGGTCCATCTACTAGAT 1353
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1354 CGTCCAGACTCGCTACAGATAGTGTGATCTCACCTGATTTATTTATCCAGTAGAA 1413
 Qy 401 AspAspValMetGluArgGlnArgLeuLeuGluSerValProAspSerValThrProLeu 420
 Db 1414 GATGATGTTATGGAAGGCGAGAGACTCATTCAGTCAGTGCCTGATTCGTGACGCCACTA 1473
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
 Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1534 CCCCAGAGTCAAGAGAGGAAATGAGTTATTTTGGCTCTGATTCGAAACACAGGTTTC 1593
 Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1594 CGTCAITTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAACGATCCAGTGT 1653
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
 Db 1654 GGGCTGCTGCTCCAAAGTATTTCAAGTGTCTCATCAAGAGGAGATAGCAATATACCAGT 1713
 Qy 501 GlyGluTrpGluValLeuGluValArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 Db 1714 GGTCAATGGGAAGTCTTGGCCGGCATGATCTATATATCCAAAGTTGATGAGTCAAGG 1773
 Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
 Db 1774 CTGGTATATTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGT 1833
 Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560

Db 1834 TAGCTAAATCTCGAGAGGTGACAAAGGCTGACTGACCTACTACACTTCTTGCTGC 1893
Qy 561 ILeSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATACGTGAGCACTGTGACTTCTTTATAGTAGTATAGTACACAGAGAGTCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCTTTACAAAGCTATCAAGTCTGAAGATGACCCAACTTGCMAAACAAGGATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGCCCAACCAATTTGGATTGAGT----- 2036
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2036 ----- 2036
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2037 -----CCTCAGGTGCAGTTG 2051
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2052 GTGAATAATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGT 2111
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2112 TATGTGTTGTAGTGATAGACACAGGGGATCTGTCCAGGGGCTTAAATTTGAAGGC 2171
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2172 GCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAAT 2231
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2232 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGCATCCAGCGCTGGTCTAT 2291
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2292 GGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTCAGATATCTTCAGGTTGCTATT 2351
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2352 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGCAAGCTTATATG 2411
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2412 GGTCACTCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2471
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAnVal 820
Db 2472 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGCTTTCTCTGGATGAGAATGC 2531
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2532 CAITTTGACATACACAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGAT 2591
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2592 TTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTAT 2651
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluLeuLeuGlySerArgIleAlaLeuLys 880
Db 2652 GAATGCTATCTTTTGCACTACCTTTCAAGAAACCTTGATCAGCTATTGCTGCTTAAAA 2711
Qy 881 ValIle 882
Db 2712 GTGATA 2717
RESULT 13
ABK83331
ID ABK83331 standard; cDNA; 4676 BP.

XX AC ABK83331;
XX 12-AUG-2002 (first entry)
XX cDNA encoding human DPRP-1 splice variant #7.
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX Homo sapiens.
OS
XX WO200231134-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US031874.
XX 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX P-PSDB; ABG61600.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 72-73; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX SQ Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;
Alignment Scores: Length: 4676
Pred. No.: 0 Matches: 831
Score: 4395.00 Conservative: 0
Percent Similarity: 94.2% Mismatches: 1
Best Local Similarity: 94.2% Indels: 51
Query Match: 6 Gaps: 1
US-10-825-632-1 (1-882) x ABK83331 (1-4676)
Qy 1 MetaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGTAGATATTGAACTCGGACTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATATTGAATCACAGGATCGCCCTAAATTTGGAGCCCTTTTATGTTGAGCGTAT 333

QY 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCCTGGAGTCAGCTTAAAGAGCTGCTGGCGATACAGAAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCCACCATGATTTTCATGTTGTGAGAGGAGATGATCCAGATGGACCTCATICA 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAACTTATTACCTTGGCCATGCTGGTGCAGACAGAGAAATACACTGTTTATTCT 513
QY 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTCCTCAAAATATCAATAGAGCAGCTCTTAATGCTCTCTTGGAGGCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
DB 574 GATCTTTTCAGGCCAACACTGGACTATGGATATGATTTCTCGAGAGAGAACTATTAGA 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAACCGCATTTGGAACAGTCGGAATTCCTTACGATTTATCACCAGAGAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAGGA 753
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 814 GATCCAAATTAATGCTGCTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnGluLeu 240
DB 874 TGGATATCTAACATCGTAACAGAGAGAGAGAGACTCATTTATGTGCACATGAGCTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 934 GCCAACATGGAGAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAGAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATCTGGCTATTTGGTGTGTCAAAGCTGAACTGAACTCCAGTGGTGT 1053
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
DB 1054 AAAATTCITAGAAATCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTTATGTTGGAACAAAGAGGGCAGATTCATTCCTGTTATCCTAAAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 1174 GCAATTCCTTAAAGTCACATTTTAAGATGTCAGAAATATGATGATGCTGAAGAGGATC 1233
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCATTTGAAGAGATTGAA 1293
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTGCCAGAGCTGGATGGACTCCTGAGGAAAAATATGCTGGTCCATCTCTACATGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1354 CGCTCCCACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1413
QY 401 AspAspValMetGluArgGlnArgMetIleGluSerValProAspSerValThrProLeu 420

DB 1414 GATGATGTTATGGAAGGCAGAGACTCATTCAGTCAGTCGCTGATTCGTGACGCCACTA 1473
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1474 ATTATCTATGAAGAAACACACACATCTGGATAAATATCCATGACATCTTTCATGTTTTT 1533
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1534 CCCAAAGTCACGAAGAGAAATGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
DB 1594 CGTCATTTATACAAATTTACATCTATTTTAAGGAAGCAAAATATAAACGATCCAGTGT 1653
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
DB 1654 GGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
DB 1714 GGTGAATGGGAAGTTCTTTGGCGGCATGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTAGTAGTCAGT 1833
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TACGTAAATCTCGAGAGGTCGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGACACTGTGACTTCTTTATAAGTAAGTATAGTAAACCAAGAGAAATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCCCTTTTACAAGCTATCAAGTCTGAAGATGACCCACTTGCATAAACAAGAAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
DB 2014 TGGGCCACCATTTTGGATTCAGCAGCTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
QY 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2074 TCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 2134 CAGCCTGGAAGAAATATCTCTACTGTGCTGTTCATATATGTTGGTGTG----- 2180
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB 2180 ----- 2180
QY 681 TyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
DB 2180 ----- 2180
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
DB 2181 -----GGTCAAAATAGAAATTCAGCATCAGGTGGAGGAGCTCCAATAT 2222
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
DB 2223 CTAGCTTCTCGATATGATTTTCATTCATTCATGATGCTGGGCATCCACGGCTGTCTCTAT 2282
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
DB 2283 GGAGGATACCTCTCCCTGATGTCATTAAATGCAGAGGTCAGATATCTTTCAGGGTGTCTATT 2342
QY 761 AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780

Db 2343 GCTGGGGCCCACTCTGTGGATCTCTATGATACAGGATACACGGAACGTTATATG 2402
QY 781 GlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAclGlu 800
Db 2403 GGTACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGNA 2462
QY 801 LysPheProSerGluProAenArgLeuLeuLeuHisGlyPheLeuAepGluAenVal 820
Db 2463 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTCCTGTGATGAGAATGTC 2522
QY 821 HisPheAlaHisThrSerIleLeuLeuLeuSerPheLeuValArgAlaGlyLysProTyrAep 840
Db 2523 CATTTTGACATACACAGTATATTACTGAGTTCCTTTAGTGAGGCTGGAAGCCATATGAT 2582
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2583 TTACAGATCTATCTCTCAGGAGACACAGACATGAAGAGTTCCTGAATCGGGAGACATTAT 2642
QY 861 GluLeuHisLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaAlaLeuLys 880
Db 2643 GAACCTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTGCTCTCTAAAA 2702
QY 881 Vallile 882
Db 2703 GTGATA 2708
RESULT 14
ACA92425
ID ACA92425 standard; DNA; 2929 BP.
XX
AC ACA92425;
XX
DT 15-JUL-2003 (first entry)
XX
DE DNA encoding human PMMM-10.
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;
KW antiinflammatory; thyromimetic; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AJA, Hawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI; 2003-430274/40.
DR P-PSDB; ABU92030.
XX
PT New human protein modification and maintenance molecules (PMMM), useful
for diagnosing, treating and preventing diseases or conditions associated

PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
PS Claim 5; Page 289; 311pp; English.
XX
CC The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ACA92416-ACA92455 encode the human PMMM polypeptides of the invention
XX
SQ Sequence 2929 BP; 864 A; 606 C; 661 G; 798 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 2929
Score: 4353.00 Matches: 824
Percent Similarity: 93.4% Conservative: 0
Best Local Similarity: 93.4% Mismatches: 0
Query Match: 92.6% Indels: 58
DB: 10 Gaps: 1
US-10-825-632-1 (1-882) x ACA92425 (1-2929)
QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAepCys 20
Db 203 ATGGCAGCAGCATGGAACAGACAGCTGGTGTGGATATTGAACTCGGACTGT 262
QY 21 GluGluAenIleGluSerGlnAepAepGluGluProPheTyrValGluArgTyr 40
Db 263 GAGGAGAAATTGAAATCAGAGATCGGCTAAATTTGGAGCTTTTATGTAGGCGGTAT 322
QY 41 SerTyrSerGlnLeuLysLeuAlaAepThrArgLysTyrHisGlyTyrMetMet 60
Db 323 TCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACCAAGAAATATCATGCTACATGATG 382
QY 61 AlaLysAlaProHisAepPheMetPheValLysArgAenAepProAepGlyProHisSer 80
Db 383 GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 442
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer 100
Db 443 GACAGAACTCTATTACCTTGGCCATGCTGTTGAGAACAGAGAAATACACTGTTTATTCT 502
QY 101 GluIleProLysThrIleAenArgAlaAlaValLeuMetLeuSerTyrPlysProLeuLeu 120
Db 503 GAAATTCACCAAACTATCAATAGACAGACAGCTCTTAATGCTCTCTTGAAGCCTCTTTTG 562
QY 121 AspLeuPheGlnAlaThrLeuAepTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 563 GATCTTTTTCAG----- 574
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAepTyrHisGlnGlySerGly 160
Db 574 ----- 574
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAepGlyGlyProGlnGly 180
Db 574 ----- 574
QY 181 PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet 200

Db 575 -----CAACACACCTTTTAAGGCCCAATCTAGTGAACACTAGTTGTCCCAACATACGGATG 628
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 629 GATCCAAATATATGCCCTGTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 688
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTrpValHisLenGluLeu 240
Db 689 TGGATATCTAACATCGTAACAGAGAGAAAGAGAGACTCACTTATGTGCACATGAGCTTA 748
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 749 GCCAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAA 808
QY 261 PheAspArgTrpSerGlyTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 809 TTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAAAGCTGAACCACTCCAGTGGTGT 868
QY 281 LysIleLeuArgIleLeuTrpGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 869 AAAATCTTAGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 928
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTrpProLysThrGlyThr 320
Db 929 ACATCCCTCATGTTGGAAACAGAGGAGGAGATTCATTCGCTTATCCTAAACAGGTACA 988
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 989 GCNAATCTTAAAGTCACCTTTAAGATGTCAGAAATATGATGCTGCTGAGGAAGGATC 1048
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1049 ATAGATGTCATAGATAAGGAATTAATCAACCTTTTCAGATCTCTATTGAAGAGTTGAA 1108
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIleLeuLeuAsp 380
Db 1109 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATATGCTGGTCCATCTACTAGAT 1168
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1169 CGTCTCCAGACTCGCTCAGATAGTGTGATCTCAGCTGAGTCTGCTGATTCGAGGAGTAA 1228
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1229 GATGATGTTATGGAAAGCAGAGACTCATTGAGTCAGTGGCTGATTCGTGAGCCACTA 1288
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1289 ATTATCTATGAAGAAACACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1348
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1349 CCCCAAAGTCAGAGAGAAATGAGTTATTTTTCCTCTGATGCAAAACAGGTTTC 1408
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTrpLysArgSerSerGly 480
Db 1409 CGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAAOCATCCAGTGT 1468
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 1469 GGCTCCCTGCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATACCAGT 1528
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1529 GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1588
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1589 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGT 1648
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560

Db 1649 TAGCTAAATCTGGAGAGGCTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1708
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTrpSerAsnGlnLysAsnProHisCys 580
Db 1709 ATCAGTTCAGCACTGTGACTTCTTTATATAAGTAAGTATAGTAACCAAGAAAGATCCCACTGT 1768
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1769 GTGTCCCTTTTACAGCTATCAAGTCTCAAGTCAACCACTTGCACAAACAAAGGATTT 1828
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTrpThrProGluIlePhe 620
Db 1829 TGGGCCACCATTTTGGATTCAGCAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 1888
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1889 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA 1948
QY 641 GlnProGlyLysLysTrpProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1949 CAGCCTGGAAAGAAATATCTTACTGTCTGTTTATATATGTTGTTCTCAGTGCAGTTG 2008
QY 661 ValAsnAsnArgPheLysGlyValLysTrpPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2009 GTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGT 2068
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2069 TATGTGTTGTAGTATAGACACACAGGGGATCTGTCAACGAGGGCTTAAATTTTGAAGGC 2128
QY 701 AlaPheLysTrpLysMetGlyGlnIleGluIleAspGlnValGluGlyLeuGlnTyr 720
Db 2129 GCCTTTAAATATAAAATGGGTCAATAGAAATGACGATCAGGTGGAGGACTCCAATAT 2188
QY 721 LeuAlaSerArgTrpAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2189 CTAGCTTCGATATGATGATTTTCACTTAGATCGTGGGCACTCCAGCGCTGTCTCTAT 2248
QY 741 GlyGlyTyrLeuSerSerMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2249 GGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTCAAGATATCTTCAGGGTTGCTATT 2308
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2309 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2368
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2369 GGTCACCCCTGACCAAGANTGAACAGGCTATTACTTAGGATCTGTGGCCATCCAGACGAA 2428
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2429 AAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC 2488
QY 821 HisPheAlaHisThrSerIleLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2489 CATTTTGCACATACCACTAGTATATTACTGAGTTTTTTAGTGGGCTGGAAAGCCATATGAT 2548
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2549 TTACAGATCTATCTCCAGGAGACACACGATTAAGATTTCTGATTCGGGAGACATATAT 2608
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2609 GAAGTCACTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA 2668
QY 881-Valle 882
Db 2669 GTGATA 2674

RESULT 15
ADV43982
ID ADV43982 standard; cDNA; 2349 BP.

```

XX ADVA3982;
XX
XX 10-MAR-2005 (first entry)
XX
XX Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 1610.
XX
XX microarray; psychoneuroendocrinimmune; chronic fatigue;
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW cancer; neoplasm; infection; expressed sequence tag; ss.
XX
XX Homo sapiens.
XX
XX WO2004108899-A2.
XX
XX 16-DEC-2004.
XX
XX 04-JUN-2004; 2004WO-US017686.
XX
XX 04-JUN-2003; 2003US-0475915P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nicholson A, Vernon SD;
XX
XX WPI; 2005-031682/03.
XX
XX New microarray comprising probes for genes involved in
XX psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
XX condition associated with PNI activity, e.g., inflammatory or infectious
XX diseases.
XX
XX Claim 1; SEQ ID NO 1610; 254pp; English.
XX
XX The invention relates to a new microarray which comprises probes for
XX genes involved in psychoneuroendocrinimmune (PNI) activity. The
XX microarray is useful in diagnosing a condition associated with PNI
XX activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
XX cancer and infection. The present sequence represents a
XX psychoneuroendocrinimmune gene expressed sequence tag. Note the
XX specificatio mentions SEQ ID NO of up to 3314 but only sequences up to
XX SEQ ID NO 1829 are provided.
XX
XX SQ Sequence 2349 BP; 723 A; 459 C; 508 G; 659 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2349
Score: 4118.00 Matches: 782
Percent Similarity: 88.7% Conservative: 0
Best Local Similarity: 88.7% Mismatches: 0
Query Match: 87.6% Indels: 100
Df: 14 Gaps: 1

US-10-825-632-1 (1-882) x ADVA3982 (1-2349)
QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 1 ATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTTGAAACTGGGACTGT 60
QY 21 GluGluAanIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 61 GAGGAGAAATATTGAATACACAGATCGGCCTAAATTTGGAGCCTTTTATGTGAGCGGTAT 120
QY 41 SerTTPSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 121 TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACCAGAAAATATCATGGCTACATGATG 180
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAanAspProAspGlyProHisSer 80
DB 181 GCTAAGGCCACCATGATTTTATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 240
QY 81 AspArgLysTyrTyrLeuAlaMetSerGlyGluAanArgGluAanThrLeuPheTyrSer 100

```

```

DB 241 GACAGAACTATTACCTTGCCATGCTGGTGAGAAACAGAAAAATACATGTTTTATTCT 300
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 301 GAAATTCCTCCAAAATCATCAATAGACAGCAGTCTTAATGCTCTCTTGGGAAGCCTCTTTTG 360
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuAsp 140
DB 361 GATCTTTTTTCAGGCACACTGAGATCTGGAATGTATTTCTCGAGAAGAGAACTATTAAAGA 420
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 421 GAAAGAAAACGCAATTGGAACTCGGAATTCCTTACGATTTATCACCAGGAATGGA 480
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 481 ACAATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGAGAGGCCACAAAGGA 540
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 541 TTTAGCCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACCGATG 600
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 601 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTCCTTTATATACATAGCAACGATATT 660
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
DB 661 TGGATATCTAATCATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCAACAATGAGCTA 720
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
DB 721 GCCAATCATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA 780
QY 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 781 TTGTGATAGATATCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACCTCCCACTGGTGGT 840
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
DB 841 AAAATCTTTAGAAATTTCTATATGAAGAAATGAAATGAACTGAGGTGGAAATTTATTCATGT 900
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 901 ACATCCCTTATGTTGGAAACCAAGGAGGCGCAGATTCATTCCGTTTATCTAAACAGGTACA 960
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 961 GCAAAATCCTAAAGTCACCTTTTAAAGATGTACAGAAATAATGATGATGCTGAAGGAAGGATC 1020
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1021 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCCTATTGGAAGGAGTTGAA 1080
QY 361 TyrIleAlaAspAlaGlyTyrThrProGluGlyLysTyrValAtpSerIleLeuLeuAsp 380
DB 1081 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAATATGTTGGTCCATCTACTAGAT 1140
QY 381 ArgSerGlnThrArgGluIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1141 CGTCCCAAGCTCGCTACAGATGTTGATCTCACCCTGAATTTATTTATCCCACTAGAA 1200
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
DB 1201 GATGATGTTATGGAAGGAGAGACTCAATGAGTCAGTGCCTGATTCGAGACCCACTA 1260
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1261 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1320
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1321 CCCCAAGTCAGGAAGAGAAATTGAGTTTATTTTGCCTCTGTAATCAAAACAGGTTC 1380

```

```
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1381 CGTCAATTTATACAAATTTACATCTATTTAAAGGAAGCAAAATATAAACGATCCAGTGT 1440
Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysGluLysGluLysIleThrSer 500
Db 1441 GGGCTGCCTGCTCCCAAGTGAATTCAGGTGCTCTCAAGAGGAGATAGCAATTTACAGT 1500
Qy 501 GlyGluTyrProGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1501 GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1560
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1561 CTGGTATATTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1620
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1621 TAGCTAAATCTCGAGAGGTGACAAGGCTGACCTGACCGTGGCTACTCACATTTCTTGTGTC 1680
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1681 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACCAAGAAATCCACACTGT 1740
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1741 GTGTCCCTTTACAGCTATCAAGTCTCGAGATGACCACTTCGAAACAAAGGAATTT 1800
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPhe 620
Db 1801 TGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 1860
Qy 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1861 TCCTTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA 1920
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1921 CAGCTGGAAAGAAATATCTACTGTGCTGTTTATATATGTTGCTCTCAG----- 1971
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1971 ----- 1971
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 1971 ----- 1971
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 1971 ----- 1971
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 1971 ----- 1971
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 1972 -----GTTGCTATT 1980
Qy 761 AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluAtgTyrMet 780
Db 1981 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2040
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2041 GGTCACTCTGACCAAGATGACAGGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAA 2100
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2101 AAGTTCCCTCTGAACCAAAATCGTTTACTGTCTCTTACATGGTTTCTCGATGAGATGTC 2160
```

```
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2161 CATTTTGCACATACCAGTATATTTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGAT 2220
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2221 TTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTAT 2280
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2281 GAACTGCATCTTTTGGCACTTACCTTCAAGAAAAACCTTGGATCAGCTATTGCTGCTTAAAA 2340
Qy 881 ValIle 882
Db 2341 GTGATA 2346
```

Search completed: April 15, 2006, 00:34:34
Job time : 1298 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 00:34:42 ; Search time 8334 Seconds
(without alignments)
4951.549 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700
Sequence: 1 MAANMETQLGVETPFTADC.....HLHYLQENLGSRIAALKVI 882

Scoring table: BLOSUM62

Ygapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Ygapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+ p2n.model -DB=US10825632/runat 14042006 092601 4694/app query.fasta_1
-Q=/abes/ABSSWEB spool/US10825632/runat 14042006 092601 4694/app query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes08
-USER=US10825632 @CGN 1.1 5315 -runat 14042006 092601 4694 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4528.5	96.4	3143	AK016546	AK016546 Mus muscu
2	4403	93.7	5517	AK029788	AK029788 Mus muscu
3	4277	91.0	2649	AY411615	AY411615 Homo sapi
4	4138	88.0	2649	AY411616	AY411616 Pan trogl
5	4116.5	87.6	2634	AY411617	AY411617 Mus muscu
6	3621	77.0	2292	CR609512	CR609512 full-leng
7	2833	60.3	3327	AK050021	AK050021 Mus muscu

8	2823	60.1	3457	4	AK078301	AK078301 Mus muscu
9	2820	60.0	3376	4	AK050023	AK050023 Mus muscu
10	2139.5	45.5	3243	4	HSN805448	AL834376 Homo sapi
11	1991.5	42.4	3128	4	CR627380	CR627380 Homo sapi
12	1744	37.1	2767	11	DO053109	DO053109 Homo sapi
13	1728	36.8	1265	4	AF176779	AF176779 Homo sapi
14	1720.5	36.6	1292	4	AF175225	AF175225 Homo sapi
15	1693.5	36.0	1042	3	BM557438	BM557438 AGNCOURT
16	1590	33.8	1013	7	CN647232	CN647232 ILLUMIGEN
17	1575	33.5	1027	7	CN646612	CN646612 ILLUMIGEN
18	1555	33.1	1278	4	AF173382	AF173382 Homo sapi
19	1554	33.1	902	8	DN517062	DN517062 1257572 M
20	1542	32.8	957	5	BQ675006	BQ675006 AGNCOURT
21	1539.5	32.8	910	5	BQ675260	BQ675260 AGNCOURT
22	1505.5	32.0	1041	2	BI084090	BI084090 602869453
23	1462	31.1	857	8	DN519771	DN519771 1262448 M
24	1457	31.0	840	8	DN514280	DN514280 1251358 M
25	1430	30.4	968	5	BQ671635	BQ671635 AGNCOURT
26	1427	30.4	910	5	BK372276	BK372276 BX372276
27	1396	29.7	914	5	BK390898	BK390898 BX390898
28	1391	29.6	1984	4	CR612208	CR612208 full-leng
29	1391	29.6	1992	4	CR603620	CR603620 full-leng
30	1370.5	29.2	853	2	BI223892	BI223892 602941035
31	1354	28.8	1082	5	BK2339476	BK2339476 603322338
32	1349	28.7	875	8	DN068337	DN068337 JGI_CABDA
33	1343	28.6	753	1	AL040398	AL040398 DKFZp434A
34	1343	28.6	766	6	CD803232	CD803232 UI-M-GVO-
35	1342	28.6	789	2	BG709118	BG709118 602675382
36	1342	28.6	3726	4	AK039652	AK039652 Mus muscu
37	1325	28.2	1090	3	BM910838	BM910838 AGNCOURT
38	1303	27.7	742	7	CN427229	CN427229 170006002
39	1297	27.6	804	7	CN525619	CN525619 UI-M-HN0-
40	1295	27.6	748	6	CD803197	CD803197 UI-M-GVO-
41	1290	27.4	855	8	CD803197	CD803197 UI-M-GVO-
42	1286	27.4	746	5	BY751026	BY751026 BY751026
43	1283.5	27.3	738	7	CN427234	CN427234 170006000
44	1281	27.3	779	8	CN783792	CN783792 UI-M-HYOP
45	1277.5	27.2	985	3	BQ068650	BQ068650 AGNCOURT

ALIGNMENTS

RESULT 1
AK016546
LOCUS
DEFINITION
Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clones:4932434P09 product:DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens], full insert sequence.
AK016546
3143 bp mRNA linear HTC 03-APR-2004
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Mus musculus (house mouse)
HTC; CAP trapper.
AK016546.1 GI:12855334
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
REFERENCE
2
Meth. Enzymol. 303, 19-44 (1999)
10349636
REFERENCE
3
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muranatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
REFERENCE
2
Meth. Enzymol. 303, 19-44 (1999)
10349636
REFERENCE
3
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muranatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda.Y., Ishikawa,T., Ozawa.K., Tanaka.T., Matsunura.S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kirai,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium,
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

S The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team,
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 440, 563-573 (2002)

6 (bases 1 to 3143)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Banno,H., Carninci,P., Fukuda,S., Furumoto,K., Furuno,M., Hanganui,T., Harada,N., Hayasaka,T., Hiramoto,K., Hirokawa,T., Horii,F., Imorani,K., Ichihashi,I., Itoh,M., Izawa,M., Kasukawa,T., Katoh,C., Kawaji.J., Kojima,Y., Konno.H., Kouda.M., Koyama,S., Kuribara.C., Matsuyma,T., Miyazaki,A., Nishi,K., Oka,K., Oyama,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Sato.R., Sakai C., Sakai,K., Sano,H., Sasabe,Y., Shibata,K., Shigematu,Y., Shinagawa.A., Shiraki,T., Sobue,Y., Suzuki.H., Tagami,M., Tagawa.A., Takabaishi.F., Tanaka.T., Tejima.Y., Toyoi,Y., Yamamura.T., Yasunishi,A., Yoshida,K., Yoshioka,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, -7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail genome-res@gsc.riken.jp]
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216
Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'- GAGAGAAGATCTCAAGACTGTCTTTTITTTTTTTCVN 3']. cDNA was prepared by using thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapted of sequence [5' GAGAGAAGATCTCGATTGAATAAATTAATCCCCCCCCCC 3'] . cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector : a modified plasmid KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI ; 3' end: BamHI. Host: DH10B.
Location/Qualifiers
 1..3143
 organism="Mus musculus"
 /mol_type="mRNA"
 strain="C57BL/6J"
 db_xref=FANTOM_DB:4932434F09"
 db_xref=taxon:10090"
 clones=4932434F09"
 sex="male"
 tissue types=testis"
 clone_lib=RIKEN full-length enriched mouse cDNA library"
 dev_stage="adult"
327..3005
note=unnamed protein product; DISEPTIDYL PEPTIDASE 8 homolog [Homo sapiens] (SPTP[Q9HEMS], evidence: FASTY, 95.7% ID, 100% length, match=2649)
putative"
codon start=1

QY	220	IleTrpIleSerAsnIleValThrArgGluGluArgLeuThrTyValHisAsnGlu	239	QY	580	CysValSerLeuTyLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu	599
Db	1014	ATTTGGATATCAAACTCGTAACAGGGAAGCGGAGGATCACATACGTACACAAATGAG	1073	Db	2094	TGTGTGTCCTCTCAAACTCTCAAGTCTCTGAGGATGAGCCCGCTTCAATAAAACAAAGGAA	2153
QY	240	LeuAlaAsnMetGluGluAspAlaAArgSerAlaGlyValAlaThrPheValLeuGlnGlu	259	QY	600	PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyThrProGluIle	619
Db	1074	CTAGCCAAACATGGAAGAGATCCAGATCAGCTGGGTGGCCACCTTTGTTCTTCAGAA	1133	Db	2154	TTTTGGGCGCACCATTTTGGATTTCAGAGGTCCTCTCTGACTACACCCCTCCAGAAAT	2213
QY	260	GluPheAspArgTySerGlyTyTrpTyCysProLysAlaGluThrThrProSerGly	279	QY	620	PheSerPheGluSerThrThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAsp	639
Db	1134	GAATTGTACAGATCTCTGGCTACTGTGTGTCTCCCAAGCGGAAGAACTCTCTAGTGT	1193	Db	2214	TTTTCTTTTGAAGATCTACTTGGATTTTACACTGTATGGAATGTTGTATAAGCCCTCATGAC	2273
QY	280	GlyLysIleLeuArgIleLeuTyGluGluAsnAspGluSerGluValGluIleHis	299	QY	640	LeuGlnProGlyLysLysTyThrValLeuPheIleTyGlyGlyProGlnValGln	659
Db	1194	GGTAATAATCTTAGAATCTCTATGAAGAAATGATGATCTCGAGGTGAGATATTATCAT	1253	Db	2274	CTACAACTGGAAAGAAATACCCCACTGTGTATTATATATATGTTGCTGCCAGGTCCAG	2333
QY	300	ValThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyProLysThrGly	319	QY	660	LeuValAsnAsnArgPheLysGlyValLysTyThrPheArgLeuAsnThrLeuAlaSerLeu	679
Db	1254	GTTACGTCCCCATGTTGGAACCAAGGAGGCGAGATTCCTTTCGTTATCCCAAAACAGGC	1313	Db	2334	CTGGTGAACTCGGTTTAAAGGAGTCAAGTATTTCGCCCTGAACACCCCTGGGCTCCCTG	2393
QY	320	ThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArg	339	QY	680	GlyTyValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu	699
Db	1314	ACAGCAAAACCAAGAGTCACTTTCAAGATGTCGAGATTTGTTGATGTCGAGAGGG	1373	Db	2394	GGTTATGTGGTTGTGTGATAGACACAGGGGATCTGTCCAGGAGACTTAATTTGAA	2453
QY	340	IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVal	359	QY	700	GlyAlaPheLysTyLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln	719
Db	1374	ATTATAGATGTCATAGATAAAGAACTGGTTCAACCTTTCGAGATTCGTTGAGGAGTT	1433	Db	2454	GGCGCCTTTAAATATGAAATGGGTCAATAGAAATCGATGATCAAGTGGAGGACTCCAG	2513
QY	360	GluTyTrpIleAlaArgAlaGlyTrpProGluGlyLysTyAlaTrpSerIleLeuLeu	379	QY	720	TyrLeuAlaSerArgTyAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSer	739
Db	1434	GAATATTTGCCAGAGCGGATGACCTCCAGAGGGGAAACATGCTGCTCCATCTACTA	1493	Db	2514	TACCTAGCATCTCAGTATGACTTTCATTGACTGGATGGGCACTCCACGGCTGGTCC	2573
QY	380	AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal	399	QY	740	TyrGlyGlyTyLysSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	759
Db	1494	GACCGTTCCAGACTCACCTGCAGATAGTTCTGATCTCCCTCGATTTATTCATCCAGTA	1553	Db	2574	TATGGTGGCTACTCTCCCTGATGGCATTAATGCAGAGGTCCGATATCTTCCGGGTGGCT	2633
QY	400	GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro	419	QY	760	IleAlaGlyAlaProValThrLeuTrpIlePheTyArgSerThrGlyTyThrGluArgTy	779
Db	1554	GAAGATGATGCCATGGACAGACAGACACTTATAGAGTCAGTTCTGACTCTGGACACCA	1613	Db	2634	ATTGCTGGGCGCCAGTCACCTGTGGATCTTCTATGATACAGGATACAGGAGCGCTAT	2693
QY	420	LeuIleIleTyGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisVal	439	QY	780	MetGlyHisProAspGlnAsnGluGlnGlyTyTrpLeuGlySerValAlaMetGlnAla	799
Db	1614	CTGATCATCTATCAAGAAACACACACATCTGATAAATATCCACGATATTTTCATGTT	1673	Db	2694	ATGGGTCAACCTGACAGAAATGAACAGGCTACTACTAGGATCTGTGGCCATCGAAGCG	2753
QY	440	PheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGly	459	QY	800	GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn	819
Db	1674	TTTCCTCAAACTCATGAGATGAATGAGTTATTTTGGCTCTGATGCAAAACAGGT	1733	Db	2754	GAGAAGTTCCCTTCAGAACCAACCGGTTACTCTCTTGGCATGGGTTCTTGGATGAGAT	2813
QY	460	PheArgHisLeuTyLysIleThrSerIleLeuLysGluSerLysTyLysArgSerSer	479	QY	820	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTy	839
Db	1734	TTTCGTCACTGTATAAATCATACATCCATTTTAAAGGAGAGCAAAATATAACGGTCCAGT	1793	Db	2814	GTTCACTTTCACACACACAGTATATTTGCTGAGTTTTTAGTGAGGCTGGAAAGCCATAT	2873
QY	480	GlyGlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThr	499	QY	840	AspLeuGlnIleTyThrProGlnGluArgHisSerIleArgValProGluSerGlyLysHis	859
Db	1794	GGTGGACTACCTGCCCAAGTATTCAGTCTCAAGTCTCATCAAGAAAGAAATAACAAATACC	1853	Db	2874	GACTTACAGATCTATCTCAGGAGAGGACACAGCATCAGAGTTCTCTGAGCTGGAGAACAC	2933
QY	500	SerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	519	QY	860	TyrGluLeuHisLeuLeuHisTyLysLeuGlnLeuAsnLeuGlySerArgIleAlaLeu	879
Db	1854	AGTGGTGAATGGGAAGTACTTGGCGGCATGATCTAATATCTGGGTTGATGAAGCCAGA	1913	Db	2934	TATGAATGCACTGTCTCACTACCTTCAGGAGAACCTTGGATCCGTTATCGTGTCTCTG	2993
QY	520	ArgLeuValTyPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyValVal	539	QY	880	LysValIle 882	
Db	1914	AAGCTGTGTACTTTGAAGGCACCAAGACTCTCTTTGGAGCATCCTCTAGTACC	1973	Db	2994	AAAGTGATA 3002	
QY	540	SerTyValAsnProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCys	559	RESULT 2			
Db	1974	AGTTATCAAAACCTGGAGAGTGGTGAGGCTGACTGACCGGTGGCTACTCACACTCTCTGC	2033	AK029788			
QY	560	CysIleSerGlnHisCysAspPheIleSerLysTyThrSerAsnGlnLysAsnProHis	579	LOCUS			
Db	2034	TGCCTCAGCGGATTTGTGACTTCTTCAAGTAAGTACAGCAACAGAGAATCCAC	2093	AK029788			
				DEFINITION			
				Mus musculus adult male testis cDNA, RIKEN full-length enriched			
				library, clone:493056C15 product:DIPEPTIDYL PEPTIDASE 8 homolog			
				[Homo sapiens], full insert sequence.			
				ACCESSION			
				AK029788.1			
				GI:26081520			
				KEYWORDS			
				HTC; CAP trapper.			

5517 bp mRNA linear HTC 03-APR-2004
AK029788
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:493056C15 product:DIPEPTIDYL PEPTIDASE 8 homolog
[Homo sapiens], full insert sequence.
AK029788
AK029788.1 GI:26081520
HTC; CAP trapper.

SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.

TITLE
JOURNAL
PUBMED
High-efficiency full-length cDNA cloning
10349636

REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
PUBMED
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitasuna, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Taka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

TITLE
JOURNAL
PUBMED
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE
JOURNAL
PUBMED
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE
JOURNAL
PUBMED
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
PUBMED
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1. .5517
/organism="Mus musculus"
/mol_type="mRNA"

FEATURES
source
1. .5517
/strain="CS7BL/6J"
/db_xref="FANTOM_DB:4930560C15"
/db_xref="taxon:10090"
/clone="4930560C15"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
2221. .4837
/note="DIPPTIDYL PEPTIDASE 8 homolog [Homo sapiens] (SPTR|Q9HBM5, evidence: FASTY, 95.7%ID, 100%length, match=2649)
putative"

misc_feature
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 5517
Score: 4403.00 Matches: 827
Percent Similarity: 95.5% Conservative: 17
Best Local Similarity: 93.6% Mismatches: 17
Query Match: 93.7% Indels: 23
DB: 4 Gaps: 2
US-10-825-632-1 (1-882) x AK029788 (1-5517)
Qy 1 MetaAlaAlaMetGluThrGluGluLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 2248 ATGCCAGCAGCAATGGAAACAGACACGCTGGGTGTCAGATATTTGAACTGCAGAGTGT 2307
Qy 21 GluGlu---AsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArg 39
Db 2308 GAGGAGGGAATGGTGAATCACAGGACGCCCAACTGGAGCCCTTTTATGTCGAACGG 2367
Qy 40 TyrSerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMet 59
Db 2368 TATTCCTGAGTCAGCTGAAAAGCTGCTTCTGATACCAAAAATACCAAGCTACATG 2427
Qy 60 MetaAlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHis 79
Db 2428 ATGGCTAAGGCGCCACATGACTTATTTGTTGTTAAAGAGGAGCGGATCCAGATGGCCCTCAC 2487
Qy 80 SerAspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 99
Db 2488 TCAGACAGGGTCTATTACCTTGCCATGTCTGGTGAGAACACAGAGAAAATACACTATTTTAT 2547
Qy 100 SerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeu 119
Db 2548 TCTGAAATCCCTAAACACCATCAACAGACAGCAGTCTTAATGCTTTCTCGAAGCCCTT 2607
Qy 120 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeu 139
Db 2608 TTGGATCTTTTTCAGCGGCACACTAGACTATGGATGTATTCTCGAGAGGAAGAGTACTC 2667
Qy 140 AtgGluArgLysArgLysGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySer 159
Db 2668 AGAAGAAAGAAAGCGCATTTGGAACTGGGGATCGCAGCTTATGATTACCCAGGAAGC 2727
Qy 160 GlyThrPheLeuPheGlnAlaGlySerGlyLysTyrHisValLysAspGlyGlyProGln 179
Db 2728 GGAACATTCTGTTCAGCTGGCAGCGCATTTACCATATCAAGATGAGAGGCCACAT 2787
Qy 180 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 199
Db 2788 GGATTTACCAACACAGACTTTGGGCCCAATTTAGTGAACACTAGTTGTCCTCCAAATATACGA 2847
Qy 200 MetAspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAsp 219
Db 2848 ATGGATCCAAATTTATGCCCTGCTGATCCAGACTGGATAGCTTTCATTTCACAGCAATGAT 2907
Qy 220 IleTrrPileSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGlu 239
Db 2908 ATTTGGATATCAACCTCTGTAACCGAGGAGCGGAGGATCATCATGTCATCAACAATGAG 2967

QY	240	LeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	259	Db	3985	ATTTGGGCGCCACATTTTGGATTGAGGTCTCTTCTGACTACACCCCTCCAGAAAT	4044
Db	2968	CTAGCCAAACATGGAGAGAGATCCAGATCAGCTGGTGTGGCCACCTTTGTTCTTCAAGAA	3027	QY	619	ePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAs	639
QY	260	GluPheAspArgTyrSerGlyTyrTyrTyrCysProLysAlaGluThrThrProSerGly	279	Db	4045	TTTTTCTTTTGAAGTACTACTCTGGATTTTACACTGTATGGAATGTTGTATAGGCTCATGA	4104
Db	3028	GAATTTGACAGATCTCTGGCTACTCTGGTGTGCTCCCAAGCGGAAAGAACTCTAGTGGT	3087	QY	639	pLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlu	659
QY	280	GlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHis	299	Db	4105	CTTACACCTGGAAAGAAATACCCACTGTGTATTATCATATATGTTGGTCCCCAGGTGCA	4164
Db	3088	CGTAAATTTCTTAGAATTTCTATCAAGAAATGATGAATCTGAGGTGGAGATTATTCAT	3147	QY	659	nLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerIle	679
QY	300	ValThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyr-ProLysThrGlu	319	Db	4165	GCTGGTGAACAACTCGTTTTAAGGAGTCAAGTATTTCCGCTCAACACCTCGCCCTCCCT	4224
Db	3148	GTTACGTCCTCCCATGTTGGAAACCAAGGAGGCGAGATTCCTTCTGTTATCCCAAAA	3202	QY	679	uGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu	699
QY	319	yThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyAr	339	Db	4225	GGGTTATGTGGTGTGGTAGACAAACAGGGGATCTCTGCCAGGAGACTTAAATTTGA	4284
Db	3203	-----CA	3204	QY	699	uGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlu	719
QY	339	gIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVa	359	Db	4285	AGCGGCTTTAAATATAAATGGGTCAAAATAGAAATCGATGATCAAGTGAAGAGACTCCA	4344
Db	3205	GATTATAGATGTAGATATAAGAACTGGTTCAACCTTCGAGATTCGTTTGGAGGAGT	3264	QY	719	nTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPse	739
QY	359	lGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuIle	379	Db	4345	GTACCTAGCATCTCATGATGACTTCTTGGATCGAGTGGGCATCCACCGCTGGTC	4404
Db	3265	TGAATATATGCCAGAGCGGATGGACTCCAGAGGAAACATCGCTGGTCCATCTACT	3324	QY	739	rTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAl	759
QY	379	uAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVa	399	Db	4405	CTATGTGGTGGTACCTCTCCCTGATGGCATTAATCCAGAGGTCCGATATCTTCGGGTGGC	4464
Db	3325	AGACCGTTCCAGACTCACCTGCAGATAGTCTCGATCTCCCTGAGTTATTCATCCAGT	3384	QY	759	aIleAlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyr	779
QY	399	lGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPr	419	Db	4465	TATTGCTGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGATACACGGAGCGCTA	4524
Db	3385	AGAAGATGATGCATGTGACACAGACAGACTTATAGACTCAGTTCTCTGACTCTGTGACAC	3444	QY	779	rMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAl	799
QY	419	oLeuIleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisVa	439	Db	4525	TATGGGTCACTGACCAAGATGAACAGGGCTACTACTTAGGATCTGTGGCATGCAAGC	4584
Db	3445	ACTGATCATCTATGAAGAAACACACAGACATCTGGATAAATATCCAGATATTTTCATGT	3504	QY	799	aGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAs	819
QY	439	lPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlu	459	Db	4585	GGAGAAGTTCCTCCAGAACCAACCGGTACTCTCTTGCATGGGTCTTGTGATGAGAA	4644
Db	3505	TTTTCTCAACTCATGAAGATGAATGAGTTTATTTTGGCTCTGATGATCAACACAGG	3564	QY	819	nValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTy	839
QY	459	yPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSe	479	Db	4645	TGTTCACTTTGCACACACCATATATTCAGTATTTTGTAGTGGGCTGGAAGCCATA	4704
Db	3565	TTTTCTGTCATCTGTATATAAATCACTCCATTTTAAAGAGAGCAATATAAAGCGTCCAG	3624	QY	839	rAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHi	859
QY	479	rGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleTh	499	Db	4705	TGACTTACAGATCTATCTCAGAGAGGCACAGCATCAGATTCCTGAGTCTGGAGNACA	4764
Db	3625	TGGTGGACTACCTGCCCCAAGTGATTTTCAAGTGTCTTATCAAGAGAAATAACAATTAC	3684	QY	859	sTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaIle	879
QY	499	rSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValAr	519	Db	4765	CTATGAACCTGACCTGCTCCACTACTCTCAGAGAACCTTGGATCGGTATCGTGTCTCT	4824
Db	3685	CAGTGTGTAATGGGAAGTACTTGGCGGCATGGATCTAATATCTGGGTTGATGAAGCCAG	3744	QY	879	uLysValIle	882
QY	519	gArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleuTyrValVa	539	Db	4825	GAAAGTGATA	4834
Db	3745	AAAGCTGGTGTACTTTGAAGGCACCAAGACTCTCTTTGGAGCATCACCTGACGTGAC	3804	RESULT 3			
QY	539	lSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCy	559	AV411615			
Db	3805	CAGTTATGCACACCTCGAGAGAGTGGTGAAGCTGACTGACCGTGGCTACTCACACTCTTG	3864	AV411615			
QY	559	sCysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHi	579	AV411615			
Db	3865	CTGCCCTGACCGGCATTTGACTTCTTCAATAGTATAGTACAGCACACGAGAAATCCACA	3924	GI:39767583			
QY	579	sCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu	599	KEYWORDS			
Db	3925	CTGTGTGCTCTTACAAACTCTCAAGTCTCTGAGGATGACCCAGTTTCATTAACAAGGA	3984	GSS			
QY	599	uPheTyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIle	619	SOURCE			
				ORGANISM			
				Homo sapiens			
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
				Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
				Hominidae; Homo.			
				REFERENCE			
				1 (bases 1 to 2649)			

2649 bp DNA linear GSS 16-DEC-2003
 AV411615
 LOCUS
 DEFINITION Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 AV411615
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2649)

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2649)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..2649
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>2649
 /gene="DPP8"
 /locus_tag="HCM4283"
gene
ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2649
 Score: 4277.00 Matches: 815
 Percent Similarity: 92.4% Conservative: 0
 Best Local Similarity: 92.4% Mismatches: 67
 Query Match: 91.0% Indels: 0
 DB: 10 Gaps: 0

US-10-825-632-1 (1-882) x AY411615 (1-2649)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
 Db 1 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGGTTGAGATATTGAAACTCGGACTGT 60
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 61 GAGGAGATATTGAAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 120
 Qy 41 SerTyrSerGlnLeuLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 121 TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACCAGAAATATCATCGCTACATGATG 180
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAspProAspGlyProHisSer 80
 Db 181 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGGACTCATTTCA 240
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 241 GACAGAACTATTACCTTGCCTCATGCTGGTGAGAACAGAGAAATACATGTTTATTCT 300
 Qy 101 GluileProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 301 GAAATTTCCCAAACTATCAATAGACGACGAGTCTTAATGCTCTCTTTGGAAGCCTCTTTT 360
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
 Db 361 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAAGA 420
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 421 GAAAGAAAACCGCATTTGGAAACAGTCGGAATTGCTTCTTACGATTATCACCAAGGAAGTGA 480
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 481 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGCTAAAGATCGAGGGCCACAAAGA 540

Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 541 TTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATG 600
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAspIle 220
 Db 601 GATCCAAATATATGCCCTGCTGATCCAGATCGATTCCTTTATATACATACACAGATAT 560
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
 Db 661 TGGATATCTAACATCGTAACAGAGAAGAAGAGACTCATTATGTGCAATAGACTA 720
 Qy 241 AlaAsnMetGluGluAspAlaIleArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
 Db 721 GCCAATCTGGAGAGATGCCAGATCAGCTGGAGTGGCTACCTTGTCTTCCAAAGAA 780
 Qy 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
 Db 781 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGT 840
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 Db 841 AAAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGT 900
 Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 901 ACATCCCTATGTTGGAAACAGGAGGCGAGATTCATTCGCTTATCTCTAAACAGGTACA 960
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlyArgIle 340
 Db 961 GCAAATCTTAAAGTCACCTTTAAGATGTCAGAAATATATGATGATGCTGAGGAAGGNN 1020
 Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1021 NNN 1080
 Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1081 NNN 1140
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1141 CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1200
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1201 GATGATGTTATGGAAAGGCAGAGACTCATTTGAGTCAGTGCCTGATCTGTGACGCCACTA 1260
 Qy 421 IleIleTyrGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1261 ATTATCTATGAAGAACACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1320
 Qy 441 ProGlnSerHisGluGluGluIleGluPheIleAlaSerGluCysLysThrGlyPhe 460
 Db 1321 CCCAAAGTCACAGAGAGAAATTTGAGTTTATTTTTCCTCTGATGCAAAACAGGTTTC 1380
 Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1381 CGTCATTTATACAAATTTACATCTATTATTAAGGAAGCAATATTAACATCCATCGTGGT 1440
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysGluLeuIleAlaIleThrSer 500
 Db 1441 GGGCTGCTGCTCCAAANN 1500
 Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 Db 1501 NNN 1560
 Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
 Db 1561 CTGGTATATTTTGAAGGCACCAAGACTCCCTCTTAGAGCATCACCTGTACGTAGTCAGT 1620

QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 DB 1621 TACGTAATCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATCTTGCTGC 1680
 QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 DB 1681 ATCAGTCAGCAGCTGTGACTCTCTTATTAAGTAGTATAGTAACCAAGAGAAATCCACACTGT 1740
 QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
 DB 1741 GTGTCCCTTTTCAAGCATATCAAGTCTCAAGATGACCCAACTTGCAGAAACAAAGGAATTT 1800
 QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
 DB 1801 TGGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTTC 1860
 QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
 DB 1861 TCTTTTGAAGTACTACTCGAATTACATTGTATGGGATGCTCTCAAGGCTCATGATCTA 1920
 QY 641 GlnProGlyLysLeuTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
 DB 1921 CAGCCTGGAAGAATATCTCTACTGCTGCTTTCATATATGCTGCTCAGGTGCGAGTTG 1980
 QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
 DB 1981 GTGAATATCGGTGTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGT 2040
 QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
 DB 2041 TATGTGTTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGC 2100
 QY 701 AlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyr 720
 DB 2101 GCCTTTAAATATAAATCGGTCAATAGAAATTCAGCATCAGGTGAAGGACTCCAATAT 2160
 QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
 DB 2161 CTAGCTTCTCGATATGATTTTCAATGATCTAGATCGTGTGGCATCCAGCGTGGCTCTAT 2220
 QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
 DB 2221 GGAGGATACCTCTCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGTTGCTATT 2280
 QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
 DB 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATG 2340
 QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
 DB 2341 CGTCACCTTGACCAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCANN 2400
 QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
 DB 2401 NNNNNNNNNNNNACCAATCGTTACTGCTCTTACATGCTTCTTCTGGATGAGAATGC 2460
 QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
 DB 2461 CATTTTGCATACACAGATATATTCTGAGTTTTTTAGTGGGGCTGGAAGCCATATGAT 2520
 QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
 DB 2521 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTGAACTCGGAGAACATTAT 2580
 QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
 DB 2581 GAACTCACTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA 2640
 QY 881 ValIle 882
 DB 2641 GTGATA 2646

RESULT 4

AY411616 2649 bp DNA linear GSS 16-DEC-2003
 LOCUS Pan troglodytes DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY411616
 VERSION AY411616.1 GI:39767584
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee).
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Pan.
 REFERENCE 1 (bases 1 to 2649)
 Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferrier, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 REFERENCE 2 (bases 1 to 2649)
 Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferrier, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 source 1..2649
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 gene <1..>2649
 /gene="DPP8"
 /locus_tag="HCM4283"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2649
 Score: 4138.00 Matches: 791
 Percent Similarity: 89.7% Conservative: 0
 Best Local Similarity: 89.7% Mismatches: 91
 Query Match: 88.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-825-632-1 (1-882) x AY411616 (1-2649)
 QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 DB 1 ATGGCAGCAGCATGGAACACAGCAGCTGGGTGTGAGATATTGAACTCGGACTGT 60
 QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 DB 61 GAGGAGATATTGAAATCACAGGATCGGCTAAATTTGGAGCCTTTTATGTTAGCGGTAT 120
 QY 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 DB 121 TCCTGGAGTCAGCTTAAAGGCTGCTGCCGATACAGAAATATCATGCTACATCATG 180
 QY 61 AlalaAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 DB 181 GCTAAGCGCACCATGATTTCATGTTGTGAAGAGGAATGATCCAGATGGCCCTCATCA 240
 QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 DB 241 GACAGATCTATTACCTTGNNNNNNGTGGAGACAGAGAAATACACTGTTTATCT 300
 QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120

Db	301	GAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATAGCTCTCTTGGAAGCCTCTTTTG	360
Qy	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
Db	361	GATCTTTTTCAGGCAACACTGGACTATGAAATGTATCTCGAAGAAGAACTATTAAAG	420
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Db	421	GAAGAAGAAACGCAATGGAAACAGTCGAATTTGCTTCTTACGATTATACCAAGGAAGTGA	480
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Db	481	ACATTTCTGTTTCAAGCTGGTAGTGGAAATTTATCAGCTAAAGATGGAGGCCACAAAG	540
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Db	541	TTTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTTGTCCCAACATACGATG	600
Qy	201	AspProLysLeuCysProAlaAspProAspTpileAlaPheIleHisSerAsnAspIle	220
Db	601	GATCCAAAATATGCCCTGCTGCATCCAGACTGGATTGCTTTTATACATAGCAACGATAIT	660
Qy	221	TpileSerAsnIleValThrArgGluArgLeuThrArgLeuThrTyrValHisAsnGluLeu	240
Db	661	TGGATATCTACATCGTAACGAGAGAAGAGGAGNCTCATTATGTGCACTAGCTA	720
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
Db	721	GCCAACTGGAAGAAGATGCCAGATCAGCTGGAGTTNCTACTTTGTTCTCCAAGAAGAN	780
Qy	261	PheAspArgTyrSerGlyTyrTptTptCysProLysAlaGluThrThrProSerGlyGly	280
Db	781	NNNNATAGATATNCTGGCTATTTGGTGGTGTCCAAAGAGCTGAACAACCTCCCAAGTGGT	840
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
Db	841	AAAATTTCTAGAAATCTATATGAGAAAAATGATGAATCTGAGGTGGAAATTTATTCATGT	900
Qy	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Db	901	ACATCCCTCATTTGGAAAACAAGAGGCGCAGATTCATTCCGTTATCTCTAAAACAGGTACA	960
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Db	961	GCAAAATNCTAAAGTCACTTTTAAGATGTCAGAAAATAATGATTGATGCTGAAGGAAGNN	1020
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Db	1021	NN	1080
Qy	361	TyrIleAlaArgAlaGlyTptThrProGluGlyLysTyrAlaTptSerIleLeuLeuAsp	380
Db	1081	NN	1140
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1141	CGCTCCAGACTCGCGTCGATAGTGTGATCTACCTGAATATATTTATCCCAAGTAGAA	1200
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1201	GATGATGTTATGGAAGGACAGACTCATTTAGTTCAGTGCCTGATCTGTGACGCCACTA	1260
Qy	421	IleIleTyrGluGluThrThrAspIleTptIleAsnIleHisAspIlePheHisValPhe	440
Db	1261	ATTANGNATGAAGAAACACAGACATCTGATATAATATCCATGAATCTCTTCATGTTTTT	1320
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Db	1321	CCCCAAGTCACGAAGAGAAATTTAGTATTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1380
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Db	1381	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATAAAACGATCCAGTGGT	1440

[illegible]

```

Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2521 TTACAGATCTATCTCTCAGGAGACACACAGTTCCTGAATCTGGAGAACATTAT 2580
Qy 861 GluLeuHisLeuHisTyrLeuGlnGluLeuGlySerArgIleAlaAlaLeuLys 880
Db 2581 GAATGTCATCTTTTGCACCTTCAAGAAACCTTGGATCAGCTATTTGCTCTAATA 2640
Qy 881 ValIle 882
Db 2641 GTGATA 2646

RESULT 5
LOCUS AY411617 2634 bp DNA linear GSS 16-DEC-2003
DEFINITION Mus musculus DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411617
VERSION AY411617.1 GI:39767585
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2634)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2634)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2634
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2634
/genes="DPP8"
/locus_tag="HCM4283"
gene
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2634
Score: 4116.50 Matches: 782
Percent Similarity: 90.4% Conservatives: 16
Best Local Similarity: 88.6% Mismatches: 78
Query Match: 87.6% Indels: 7
DB: 10 Gaps: 2

US-10-825-632-1 (1-882) x AY411617 (1-2634)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 1 ATGCGCAGCAATGGAACAGACAGCTGGGTGCGAGATATTGAAACTGCAGAGTGT 60
Qy 21 GluGlu---AsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArg 39
Db 61 GAGAGGAAATGGTGAATCAGAGGACCGACCAACTGGAGCCCTTTTATGTGCAACGG 120
Qy 40 TyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMet 59

```

```

Db 121 TATTTCTGGAGTCAGCTGAAAAAGCTGCTTGTGATACCAGAAATACCACGGCTACATG 180
Qy 60 MetAlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHis 79
Db 181 ATGGCTAAGGGCGCCACATGACCTTATGTTGTAAAGAGCGAGATCCAGATGGCCCTCAC 240
Qy 80 SerAspArgIleTyrTrpLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 99
Db 241 TCAGACAGGGTCTATTACCTTGCCTGCTGTGTGAGAAACAGAGAAAATACACTATTTTAT 300
Qy 100 SerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeu 119
Db 301 TCTGAATTCCTTAAACCATCAACAGCAGCTCTTAATGCTTTCTCGAGACCCCTT 360
Qy 120 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeu 139
Db 361 TTGGATCTTTTTCAGGCGACACTAGACTATGGATGTATTCTCGAGAGGAAGACTACTC 420
Qy 140 ArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySer 159
Db 421 AGAGAAAGAAAGCGCAATTGGAACCGTGGGGATCGCAGCTTATGATTACCCACCGAAGC 480
Qy 160 GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGln 179
Db 481 GGAACATTTCTGTTCAAGCTGGCAGCGCATTTACCATATCAAGATGGAGGACCAT 540
Qy 180 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 199
Db 541 GGATTTACACACACNGCCTTTGGGCCCAATTTAGTGGAACTAGTTGTCCCAATATACGA 600
Qy 200 MetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAsp 219
Db 601 ATGGATCCAAAATTATGCTCTGCTATCCAGACTGGATAGCTTTCAITTCACAGCAATGAT 660
Qy 220 IleTrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGlu 239
Db 661 ATTGTGATATCAAACTCTGTAACCGAGGAAGAGCGGAGGATCACTAGTACACAAATGAG 720
Qy 240 LeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 259
Db 721 CTAGCCCAACATGGAAGAGGATCCAGATCAGCTGGTGTGGCACTTTGTTCTTCAAGAA 780
Qy 260 GluPheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGly 279
Db 781 GAATTTGACAGATACTCTGGTACTTGGTGTGTCCTCCCAAGCGAAAGAACTCTCTAGTGT 840
Qy 280 GlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHis 299
Db 841 GGTAAATTTCTTAGAATTTCTTATGAAGAAAATGATGAATCTGAGGTGGAGATTATTCAT 900
Qy 300 ValThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGly 319
Db 901 GTTACGTCCCCCATGTTGAAAACAGAGGCGGAGATTCCTTCTGTTATCCCAAAACAGGC 960
Qy 320 ThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArg 339
Db 961 ACAGCAAAACCAAGGTCACTTTCAAGATGTCGGAGATTGTTGTGTGCTGCAGGAGGG 1020
Qy 340 IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVal 359
Db 1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080
Qy 360 GluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeu 379
Db 1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140
Qy 380 AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal 399
Db 1141 GACCGTTCCAGACTCACCTGCAGATAGTTCTCGATCTCCCTGAGTTATTTCATCCAGTA 1200
Qy 400 GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro 419

```

Db	1201	GAAGATGATGCCATGGACAGACAGAGACTTTATAGATCAGTCTCTGACCTCTGTGTGACCA	1260
Qy	420	LeuIleIleTyrGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisVal	439
Db	1261	CTGATCATCTATGAAGAAACACACAGACATCTGGATAAATATCCAGATATTTTTCATGTT	1320
Qy	440	PheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysIleThrGly	459
Db	1321	TTTCTCTCAACTCATGAAGATGAATTTGAGTTTATTTTTCCTCTGAAATGCAAAACAGGT	1380
Qy	460	PheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysIleThrLysArgSer	479
Db	1381	TTTCTCTCATCTGATATAAATCACATCCATTTTAAAGAGAGCAATAATATAAAGGTCAGT	1440
Qy	480	GlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThr	499
Db	1441	GTTGAGTACCTGCTGCCCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1500
Qy	500	SerGlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	519
Db	1501	NN	1560
Qy	520	ArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValVal	539
Db	1561	AAGCTGGTGATCTTTGAAGGACCAAGACTCTCTTTGGAGCATCCTCTGATGACCTGACC	1620
Qy	540	SerTyrValAsnProGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys	559
Db	1621	AGTTATGCAAAACCTGGAGAAGTGTGAGGCTGACTGACCGTGGCTACTCACACTCTGTC	1680
Qy	560	CysIleSerGlnHisCysAspPhePheIleSerIleTyrSerAsnGlnLysAsnProHis	579
Db	1681	TGCTCAGCGGCATTTGTGACTCTCTCATAGTAAGTACAGCAACAGCAAGATCCACAC	1740
Qy	580	CysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu	599
Db	1741	TGTGTGTCTCTCAAAACTCTCAAGTCTGAGGATGACCCAGTTCATAAAACAAAGGAA	1800
Qy	600	PheTyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIle	619
Db	1801	TTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTCTGACTACACCCCTCCAGAAAT	1860
Qy	620	PheSerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAsp	639
Db	1861	TTTTCTTTTGAAGTACTTGTGATTTTACACTGTATGGAATGTTGTATAAGCCTCATGAC	1920
Qy	640	LeuGlnProGlyLysIleTyrProThrValLeuPheIleTyrGlyGlyProGlnValGln	659
Db	1921	CTACACCTTGGAAAGAAATACCCACTGTGTTATTCATATATGTTGTTCCCCAGGTGCAG	1980
Qy	660	LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeu	679
Db	1981	CTGTTGAACATCGGTTTAAAGAGTCAAGTATTTTCGCTGCAACCCCTGGCTCCCTG	2040
Qy	680	GlyTyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGlu	699
Db	2041	GGTTATGTGTGTGTGTATAGAACACAGGGGCTCTGTACAGGAGCTTAATTTGAA	2100
Qy	700	GlyAlaPheLysTyrLysMetGlyGlnIleLeuLeuAspGlnValGluGlyLeuGln	719
Db	2101	GGCGCTTTAAATATAAATATGAGTCAATAGAAATCGATATCAAGTGGAGGACTCCAG	2160
Qy	720	TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSer	739
Db	2161	TACTATGACATCTCAGTATGACTTCACTTGGATGAGTGGGATCCACGCTGTGCTC	2220
Qy	740	TyrGlyGlyTyrLysSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	759
Db	2221	TATGGTGGCTTACCTCTCCCTGATGGCATTAATGCAAGGTCGGATATCTTTCCGGGTGGCT	2280
Qy	760	IleAlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyr	779
Db	2281	ATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAGCGCTAT	2340

Qy	780	MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla	799
Db	2341	ATGGGTACCTCTGACCAAGATGAACAGGCTACTACTTAGGATCTGTGSCCATGCAAGCA	2400
Qy	800	GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn	819
Db	2401	-----CCAAACCGGTACTCTCTTTCATGGGTCTTGGATGAGAT	2442
Qy	820	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr	839
Db	2443	GTTCACTTTGCACACACAGTATATCTGAGTTTTTAGTGAGGCTCGAAGCCATAT	2502
Qy	840	AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis	859
Db	2503	GACTTACAGATCTATCTCTCAGGAGAGCAGCATCAGAGTTCTTGAGTCTGAGAACAC	2562
Qy	860	TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeu	879
Db	2563	TATGAATGCACCTGCTCTCACTTACAGGAACCTTGGATCGGTATCGCTGCTCTG	2622
Qy	880	LysValIle 882	
Db	2623	AAAGTGATA 2631	
RESULT 6			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			

QY 806 ProAenArgLeuLeuLeuLeuHisGlyPheLeuAspGluAenValHisPheAlaHisThr 825
 Db 1868 CCAATCGTTTACTGCTTTACATGGTTTCTGGATGAGAAATGTCATTTTGCATATCC 1927
 QY 826 SerIleLeuLeuSerPheLeuValArgAlaGlyIleProTyrAspLeuGluIleTyPro 845
 Db 1928 AGTATATACAGAGTTTATAGTGGGCTGGAAAGCCATATGATTTA----- 1975
 QY 846 GlnGluArgHisSerIleArgValProGluSerGlyGluHisTyGluLeuHisLeuLeu 865
 Db 1976 CAGAGAGACACACATAGAGTTCTCGAATCGGAGAACATATATGAATGCATCTTTTG 2035
 QY 866 HisTyIleuGlnGluAenLeuGlySerArgIleAlaLeuLeuValIle 882
 Db 2036 CACTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAAAGTGATA 2086

RESULT 7
 AK050021
 LOCUS
 DEFINITION
 Mus musculus adult male liver tumor cDNA, RIKEN full-length
 enriched library, clone: C730003D12 product: DIPEPTIDYL PEPTIDASE 9
 homolog [Homo sapiens], full insert sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AK050021.1 GI:26340743
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS

1 Carninci, P., and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 10349636

2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 11042159

3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kusunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 11076861

4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

6 (bases 1 to 3327)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

TITLE
JOURNAL

COMMENT

FEATURES
Source

1 3327
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM:DB:C730003D12"
 /db_xref="taxon:10090"
 /clone="C730003D12"
 /sex="male"
 /cell_type="tumor"
 /tissue_type="liver"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 176..2764
 /note="unnamed protein product; DIPEPTIDYL PEPTIDASE 9
 homolog [Homo sapiens] (SPTRIAL47179, evidence: FASTY,
 92%ID, 99.6%length, match=2580)
 putative"
 /codon_start=1
 /protein_id="BAC34034.1"
 /db_xref="GI:26340744"
 /translations="MCSGVSPVSESHRLYVLGMPYSGRENSLLYSIPKVKREALLLL
 SWKMLDHFQFQVQKPSDSESHRLYVLGMPYSGRENSLLYSIPKVKREALLLL
 GLVSKAPDFQFVQKPSDSESHRLYVLGMPYSGRENSLLYSIPKVKREALLLL
 SKWMLDHFQFQVQKPSDSESHRLYVLGMPYSGRENSLLYSIPKVKREALLLL
 CRDQKGNFVPMKPLEIKTQCSGRMDPKICPADPAFFSINNSDLFWANSETGEE
 RLRTFCHQGSAGVLDNPKSAGVATFVIOEEDRFTGCMWCTPAAWEGSGELTKRLY
 EEVDESEVIVHVPSPALREKRTDSVYPTGSKNPKIALKLAELQTHQKQIVSCE
 KEAVQPESSLPKVEYLARAGWTRDGYAWMFLDPQRLQLVLLPPLFPAVESE
 AQRQAARAVPKVQPVLYEETVNVINVDIHPFPAQEQDQFCFLRANECKTF
 CHLYRVETLKTQDWTPLSPTEDEFEKPIKEVALTSCEMSQSPDMFVSHVSUS
 TKLYVQCTKDTPLEHLIVVSYESAGEIVRLTTLGFSHSCMSQSPDMFVSHVSUS
 TPCVHYLUSGDDDFLHKQRFWASMEANCPDPYPPPEIFTHFRADYVQLXCM1
 YKPHLQGRKHPTLVFVYGGQVQLVNNNSFKGIKYLRLNTLASLGIAVVIDGKSC
 QRGHFEGALKNQMGVEIEDQVEGQIVAEKYGIFIDLSRVAIHGSWGGFLMGLLI
 HKQVFKVALAGAPVTVMAYDITGYTERMDVFNQNSQSYGSEVALVEKLPKPNFR
 LLHLHGLDENVHFHTNVLVSQLIRAGPKYQLQIYPNERHSIRCSGESGEYEVETLLH
 FLOEHL"

CDS

ORIGIN

Alignment Scores:
 Pred. No.: 6.82e-313 Length: 3327
 Score: 2833.00 Matches: 511
 Percent Similarity: 76.8% Conservative: 134
 Best Local Similarity: 60.8% Mismatches: 193
 Query Match: 60.3% Indels: 2
 DB: 4 Gaps: 2

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akashira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M., and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in RIKEN
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissue was provided by William A. Held, Roswell Park Cancer
 Institute, Department of Molecular and Cellular Biology, Elm and
 Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully
 acknowledge.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.
 Location/Qualifiers

QY	753	SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTyrIlePheTyrAsp	772
Db	2402	CCACAAGTGTTCAGAGTAGGCATTCGGGGCGCTTCCTGTGCACGTGTGGATGGCCCTATGAC	2461
QY	773	ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu	792
Db	2462	ACAGGGTACACGGAACATACATGATGTCCCGAAATAAACACACAGGCTATGAGGCA	2521
QY	793	GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu	812
Db	2522	GGGTCTGTAGCCCTGCATGTGGAGAACTGCCCAATGAGCCTAAACCGCTGCTTATCCTC	2581
QY	813	HisGlyPheLeuAspGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeu	832
Db	2582	CACGGCTTCCTGGACGAGAACGTTTCATCTTCCACACAAATTTCTCGTGTGCCAGCTG	2641
QY	833	ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg	852
Db	2642	ATCCGAGCAGGAAAGCCATACACAGCTTCAGATCTACCCAAACGAGACATAGCATCGC	2701
QY	853	ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeu	872
Db	2702	TCCCGGAGTCCGGAGAGCATACGAGGTGACCTGCTGCACCTTTCGAGGAACACCTG	2761
RESULT 8			
AK078301		3457 bp mRNA linear HTC 03-APR-2004	
LOCUS		Mus musculus adult male olfactory brain cDNA, RIKEN full-length	
DEFINITION		homolog [Homo sapiens], full insert sequence.	
ACCESSION		AK078301.1 GI:26347124	
VERSION		HTC; CAP trapper.	
KEYWORDS		Mus musculus (house mouse)	
SOURCE		Mus musculus	
ORGANISM		Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE		1 Carninci, P., and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 10349636	
AUTHORS		2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 11042159	
TITLE		3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Owa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 11076861	
JOURNAL		4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	
PUBLISHED		5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)	
AUTHORS			

REFERENCE		6 (bases 1 to 3457) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Haeishizume, W., Hayashida, K., Hayase, N., Hiramoto, K., Hiraoaka, T., Hirokawa, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsumoto, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akanishi, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 200-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/	
FEATURES		1..3457 Location/Qualifiers	
source		organism="Mus musculus" mol_type="mRNA" /str_id="C57BL/6J" /db_xref="FANTOM_DB:6430584G11" /db_xref="taxon:10090" /clone="6430584G11" /sex="male" /tissue_type="olfactory brain" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 182..2770	
CDS		/note="unnamed protein product; DIPEPTIDYL PEPTIDASE 9 92%ID, 99.6%length, match=2580" homolog [Homo sapiens] (SPTK AAL47179, evidence: FASTY, putative" /codon_start=1 /protein_id="BAC37211.1" /db_xref="GI:26347125" /translation="MCSGVSEVQVAAAGMDMTAARFCVQKGNWDGLRSLIHGSRKSS GLVSKAPHDFQVKPDGSGHSHRLYLGLMPYSGRENSLLYSEIIPKVRKEALLLL SWKQLDHFQATPHHGVYREELRLERKRLGVGITSYDFHSESLGFLFQASNLFLH CDGKNGFMVSPMKPLEIKTCGSPMDPKICPADPAFFSFINNSDLVWVNIETGEE RRLTFCHGGSAGVLNPKSAGVATFVIOEEFDRFTGWCWPTASWEGSLKTLRLY EVDSEVEVTHVPSPALERKTDTSYRPTGSKNPKIALKLAELQTHQGLTVSECE KELVQFSSLPFKVEYIARAGWTRDGYAMFLYRPOORLQVLPLPALFIPAVESE ACORQAARAVPKNVQPFVIVVEVTNVNINHDIHPFPQAGQQDFCFRANECKTFQ CHLYVVTVELTKTDYDTEPLSPTEDEFKCPKEKEVALTSGEWELSRSGSKTWNEQ TKLVVYFQTKTLPLEHLVYVSYESAGEI VRLATLGFSSHCSMSQSFDMFVSHYSVS YKPHPTQPKRHPTVLFVYGGPQVLVNSFKGIKYLRLNTLASLGVAVVVDGSGC ORGLHEGALKNQMGVEIDQVEGLQVVAEKYGFIDLSRVAILHSGVSGVGLMGLI HKPOVKVIAIAGAPVTVMAYDTCYTERYMDPENNKGVEAGSVALHVKELMELNLR LILALGELDENVHFHTNLTSLQIRAGKPYQLIYPNERHSIRCRESGHEVETLLH FLQEHL"	
ORIGIN		Alignment Scores: 1,02e-311 Length: 3457 Pred. No.: 2823.00 Matches: 510 Score: 76.1% Conservative: 134 Percent Similarity: 60.7% Mismatches: 194 Best Local Similarity: 60.1% Indels: 2 Query Match:	

```
DB: 4 Gaps: 2
US-10-825-632-1 (1-882) x AK078301 (1-3457)
QY 35 PheTyrValGluArgTyrSerTrpSerGlnLeuIleValLeuAlaAspThrArgLys 54
DB 248 TTCTGTGTCAGAGACACTCGTGGAGTGGCTCGGTAGCATATATCCACGGCAGTGCAG 307
QY 55 TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValIysArgAsnAsp 74
DB 308 TCCTGGGCGCTCATGTTCAGCAGAGCCGCCACGACTTCAGTTGTGTCAGAGCCTGAC 367
QY 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
DB 368 GAGTCTGGCGCCCACTCTCACCGTCTCTATTACTCGGAATGCTTTACGGCAGCGGTGAG 427
QY 95 AsnThrLeuPheTyrSerGluIleProIysThrIleAsnArgAlaAlaValLeuMetLeu 114
DB 428 AACTCCCTCTCTACTCTCGAGATCCCAAGAAAGTGGGAAGGAGGCGCTGCTGCTGTG 487
QY 115 SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
DB 488 TCCTGAGACGAGTCTGGACCACTTCAGGCCACACCCACCATGGTGTCTACTCCCGA 547
QY 135 GluGluGluLeuLeuArgGluArgIleGlyThrValGlyIleAlaSerTyrAsp 154
DB 548 GAGGAGGAGTACTTCGGGAGCGCAAGCGCTGGGGCTCTTCGGAATCACTCTTATGAC 607
QY 155 TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys 174
DB 608 TTCCACAGTGAGCGGCGCTCTCTCTCCAGGCCAGCAATAGCCTGTTCACCTGACG 667
QY 175 AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer 194
DB 668 GATGTGGCAGAAATGGCTTTATGGTGTCCCGATGAGCCACTGGAGATCAAGACTCAG 727
QY 195 CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe 214
DB 728 TGTTCTGGGCCACGATGAGACCCCAAAATCTGCCCGCAGACCCCTGCTTTTCCCTTC 787
QY 215 IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluArgIleThr 234
DB 788 ATCAACAACAGTGATCTGTGGTGGCAACATCGAGACTGGGAGGAACGGCGGTCAAC 847
QY 235 TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr 254
DB 848 TTCTGTACACAGGTTCAGCTGTCTGTGACATCCCAATCAGCAGCGGTGGCCACC 907
QY 255 PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGlu 274
DB 908 TTGTTCATCCAGGAGGAGTTCGACCGCTTCACCTGGGTGCTGGTGGTGGCCCGCTCT 967
QY 275 ThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer 293
DB 968 TGGGAAGGCTCCGAAGGTCTCAAGACGCTCGCATCTCTATATGAGGAAGTGGACGAGTCT 1027
QY 294 GluValGluIleIleHisValThrSerProMetLeuGluThrArgAlaAspSerPhe 313
DB 1028 GAAGTGGAGGTTCATTTCATGTGCTCCCTCCCGCTCGGAGGAGGAAGACGGACTCCTAC 1087
QY 314 ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMet 333
DB 1088 CGCTACCCACGACAGGACGAGCAAGAACCCCAAGATTGCCCTGAAGTGGCTGAGCTCCAG 1147
QY 334 IleAspAlaGluGlyArgIleAlaAspValIleAspLysGluIleLeuGlnProPheGlu 353
DB 1148 ACGGACCATCATCGGCAAAATCGTGTCAAGCTGCGAGAAGAACTGTGTACAGCCATTACG 1207
QY 354 IleLeuPheGluGlyValGluTyrIleAlaArgIleGlyTrpThrProGluGlyIleTyr 373
DB 1208 TCCCTTTTCCCAAGTGGAGTACATCGCCGGGCTGGCTGGACACGGGACGCAAAATAT 1267
QY 374 AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerPro 393
DB 1268 GCTGGGCATATTTCTGTACCGTCCCGACGACGGCTTCAGCTTGCTTCCTACCCCT 1327
QY 394 GluLeuPheIleProValIleGluAspValMetGluArgGlnArgLeuIleGluSerVal 413
DB 1328 GCTCTCTTCATCCCGCGTGTGAGAGTGAGGCCCGCAGCGGAGGAGCTGCCAGAGCGCTC 1387
QY 414 ProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIle 433
DB 1388 CCCAAGAATGTGCAGCCCTTTGTATCATATATGAAGAATCAACCAATGTGTGATCAACGT 1447
QY 434 HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
DB 1448 CACGACATCTTCACCGCTTCTCAGCTGAGGCCGACGAGGACTTTTGTTCCTTCCTCGT 1507
QY 453 AlaSerGluCysLysThrGlyPheArgHisLeuTyrIleLysIleThrSerIleLeuLysGlu 472
DB 1508 GCCACGATCAAGACTGGCTTCTGCCACCTGTACAGGGTTCACAGTGAATTAACAC 1567
QY 473 SerLysTyrLysArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIle 492
DB 1568 AAGGACTATGACTGGACGAAACCCCTCAGCCCTACAGAAGATGAGTTTAAAGTGCCTCATC 1627
QY 493 LysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn 512
DB 1628 AAGGAGGAGTGGCTTCCAGCTGAGCGGAGTGGAGGTCTTGTGCGAGCATGGCTCCAG 1687
QY 513 IleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeu 532
DB 1688 ATCTGGGTCAACGACGACGACGAGCTGGTGTACTTCAAGGTACAAAGGACACACCGCTG 1747
QY 533 GluHisLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp 552
DB 1748 GAACATCACTCTATGTGTGTCAGCTTACGATCAGCAGCGAGATCGTGCAGCTCACCACG 1807
QY 553 ArgGlyTyrSerHisSerCysLysSerGlnHisCysAspPhePheIleSerLysTyr 572
DB 1808 CTGGCTTCTCCACAGCTGCTCCATGAGCCAGACGCTTCGACATGTTGCTGAGTCACTAC 1867
QY 573 SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp 592
DB 1868 AGCAGTGTGAGCAGCGCCCTCTGTGTATGTGTACATGTGTACAGCTGAGCGGCCCGATGAC 1927
QY 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro 612
DB 1928 CCACTGCACAAACCAACGCTTCTGGCCAGCATGTATGAGGAGGAGCAATTTGCCCCCA 1987
QY 613 AspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
DB 1988 GACTATGTGCCCTCGAGATCTTCCACTTCCACACCGCTGAGAGCGTGCAGCTCTACGGC 2047
QY 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIle 652
DB 2048 ATGATCTACAGCCACACACCTCGAACCTGGAGGAGAGCACCCACCTGTGTCTTTGTC 2107
QY 653 TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArg 672
DB 2108 TATGGGGGCCACAGGTGTCAGTTGTGTGAACAACTCTTTAAGGCGCATCAATACCTCGG 2167
QY 673 LeuAsnThrLeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCys 692
DB 2168 CTAAATACACTGGCATCTCTGGGCTATGCTGTGGTGGTGTATCGATGTTGCGGGCTCTCTGT 2227
QY 693 HisArgGlyLeuLysPheGluGluValaPheLysTyrLysMetGlyGlnIleGluIleAsp 712
DB 2228 CAGCGGGGCTGCACTTCGAGGGGCGCTGAAATAATCAATATGGGCCAGGTGAGATTGAG 2287
QY 713 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 732
DB 2288 GACCAGTGGAGGCTTGGCAGTACCTGGCTGAGAAGTATGCTTTCATTGCTTGCAGCGCA 2347
QY 733 ValGlyIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 752
```


QY 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
Db 421 GAGTCTGGCCGCCCACTCTACACCGTCTCTATTACCTCGAATGCCCTTACGCGCGGTGAG 480
QY 95 AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValIleMetLeu 114
Db 481 AACTCCCTCTCTACTCCGAGATCCCAAGAAAGTCCGGAAGGAGGCCCTCTGCTGCTG 540
QY 115 SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
Db 541 TCCTGGAGAGAGATGCTGGACCACTTCAGGCCACACCCACCACCTGTTCTACTCCGA 600
QY 135 GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAs 154
Db 601 GAGGAGAGCTACCTCGCGGAGCGCAAGCGCTGGCGCTCTTCGGAATCACTCTTATGA 660
QY 154 pTyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLy 174
Db 661 CTTCACAGTGTAGAGCGGCTCTCTCTCTTCAGGCCAGCAATAGCTGTTCCACTGCAG 720
QY 174 sAspGlyGlyProGlnGlyPheThrGlnProLeuArgProAsnLeuValGluThrSe 194
Db 721 GATGTGTGGCAAGATGGCTTTATGTCTCCCGATGAAGCCACTGGAGATCAGACTCA 780
QY 194 rCysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPh 214
Db 781 GTGTCTGGGCCACGATGACCCCAAAATCTGCCCGCAGACCTGCTCTTTTCCTT 840
QY 214 eIleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgArgLeuTh 234
Db 841 CATCAACAACAGTATCTGTGGTGGCAACATCGAGACTGGGGAGGAACCGCGGCTCAC 900
QY 234 rTyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaTh 254
Db 901 CTCTGTGCACGAGGTTCAAGTGTGTGGCAATCCCAATCAGCAGGGGTGGCCAC 960
QY 254 rPheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGl 274
Db 961 CTTTGTATCCAGGAGGATTCGACCGCTTCACTGGGTGCTGTGGTGGCCCGGCTC 1020
QY 274 uThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAsnAspGluSe 293
Db 1021 TTGGAGAGGCTCCCAAGGTCCTCAAGACGCTGCGCATCTATATGAGGAAGTGGACGATC 1080
QY 293 rGluValGluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPh 313
Db 1081 TGAAGTGGAGGTCTATCATGTGCGCTCCCGCCCTGGAGGAGAGAGACGAGCTCTTA 1140
QY 313 eArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMe 333
Db 1141 CCGCTACCCAGGACAGCGAGCAAGAACCCCAAGATGCTCCGTAAGCTGGCTGAGCTCA 1200
QY 333 tIleAspAlaGluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGl 353
Db 1201 GACGGACCATCAGGCGCAAAATCGTCAAGCTCGGAGAAAGAACTGGTACAGCCATTGAG 1260
QY 353 uIleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTy 373
Db 1261 CTCCTCTTTCCCAAGTGGAGTACATCGCCCGGGTGGTGGACACGGGCGCAAAATA 1320
QY 373 rAlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValIleLeuSerPr 393
Db 1321 TGCCTGGGCCAATGTTCTGGACCGTCCCGAGCAAGCGCTTCAGCTTGTCTCTGCCCCC 1380
QY 393 oGluLeuPheIleProValGluAspAspValMetGluArgGlnArgLeuIleGluSerVa 413
Db 1381 TGCTCTCTCTATCCCGCGCTTGAGAGTGAGGCGCCAGCGCAGGCGCTGCCAGAGCGGT 1440
QY 413 lProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIl 433
Db 1441 CCCCAGAAATGTGCAGGCGCTTGTCTATGATGAAGTCAACCAATGTCTGGATCAAGGT 1500
QY 433 eHisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePh 452

Db 1501 CCAGACATCTTCCACCCGTTTCTCAGGCTGAGGCCAGCAGGACTTTTGTTCCTCTCG 1560
QY 452 eAlaSerGluCysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGl 472
Db 1561 TGCCCAACGAATGCAAGACTGGCTTCTGCCACCTGTACAGGTCACAGTGAACCTTAAAC 1620
QY 472 uSerLysTyrLysArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIl 492
Db 1621 CAAGGACTATGACTGGAGCGAACCCTTCAGCCCTCAGAGAAGATGATTTTAAAGTGCCCAT 1680
QY 492 eIysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAs 512
Db 1681 CAAGGAGAGGTGCGCTTACCAGTGGCGAGTGGAGGTCCTGTGAGGCATGGCTCCAA 1740
QY 512 nIleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLe 532
Db 1741 GATCTGGGTCAACGAGCAGCAGAGCTGGTGTACTTTCAAGGTACAAAGACACACCGCT 1800
QY 532 uGluHisLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAs 552
Db 1801 GGAACATCACCTCTATGTGTGTCAGCTACGAGTCAGCAGCGCAGATCGTGGGCTCACCA 1860
QY 552 pArgGlyTyrSerHisSerCysLysIleSerGlnHisCysAspPhePheIleSerLysTy 572
Db 1861 GCTCGGCTTCTCCACAGCTGCTCCATGAGCCAGAGCTTCGACATGTTGTGAGTCACATA 1920
QY 572 rSerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAs 592
Db 1921 CAGCAGTGTGAGCAGCGCCACCTGTGTACATGTGTACAGCTGAGCGGCGCCGATGATCA 1980
QY 592 pProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPr 612
Db 1981 CCCACTGCACAAGCAACACACGCTTCTGGGCCAGCATGATGGAGGAGCAATTTGCCCC 2040
QY 612 oAspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGl 632
Db 2041 AGACTATGTGCGCCCTGAGATCTTCCACTTCCACACCCGTCAGAGCGTGAGCTCTACCG 2100
QY 632 yMetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrTrpThrValLeuPheIl 652
Db 2101 CATGATGTACAGGCCACACACCTTGCACCTGGAGGAGGAGCACCACCTGCTCTTCTGT 2160
QY 652 eTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheAr 672
Db 2161 CTATGGGGGCCACAGGTGTCAGTGTGTGAACACTCTCTTAAAGGTCATCAATACCTGCG 2220
QY 672 gLeuAsnThrLeuAlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCy 692
Db 2221 GCTAAATACACTGGCATCTTGGGCTATGCTGTGTGTGTCATCGATGGTGGGCTCTCTG 2280
QY 692 shiArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAs 712
Db 2281 TCAGCGGGGCGCTGCACCTTCGAGGGGGCGCTGAAAAAATAAATGGGCGAGTGAGATGA 2340
QY 712 pAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspAr 732
Db 2341 GACACAGGTGGAGGCTTGCAGTACGTGGCTGAGAAATGAGCTTTCATTGACTTGAGCGG 2400
QY 732 gValGlyIleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnAr 752
Db 2401 AGTCGCCATTCATGGCTGCTCCTACGCGCGCTCTCTCTCATCTCATGGGCTCATCACAA 2460
QY 752 gSerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAs 772
Db 2461 GCCACAAGTGTTCAGGTAGCCATTGGCGGCGCTCTCTCTCATCTCATGGGCTCATCACAA 2520
QY 772 pThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTrLe 792
Db 2521 CACAGGTACACGGAACGATACATGGATGTCGCCGAAAAATAACACGACGAGCTATGAGGC 2580
QY 792 uGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLe 812


```

Qy 477 gSerSerGlyGlyLeuProAlaProSerAspPheIysCysProIleIysGluGluIleAl 497
Db 780 GAGTGAAGCCCTTCAGCCCGGGGAGAGATGAATTAAAGTGCCCAATTAAAGAGAGATGC 839
Qy 497 aIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGI 517
Db 840 TCTGACCAAGCGGTGATGGAGGTTTGGCGAGGCGACGGCTCCAAGATCTGGGTCAATGA 899
Qy 517 uVala-gArgLeuValTrpPheGluGlyThrIysAspSerProLeuGluHisIleLeuTy 537
Db 900 GGAGACCAAGCTGGTACTTCCAGGCGACCAAGACACAGCGCTGGAGACCAACCTCTA 959
Qy 537 rValValSerTrpValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHi 557
Db 960 CGTGGTTCAGTATGAGCGCGCGCGAGATCGTACGCTTACACCGCGCGCTCTCCCA 1019
Qy 557 sSerCysAspIleSerGlnHisCysAspPheIleSerIysTyrSerAsnGlnLysAs 577
Db 1020 TAGTGTCTCCATGAGCCAGAACTTCGACATGTTCTGTCAGCCACTACAGCAGCGTGAC 1079
Qy 577 nProHisCysValSerLeuTrpIysLeuSerSerProGluAspAspProThrCysIysTh 597
Db 1080 GCGCCCTGCGTGCAGCTTACAAGCTGAGCGCGCGCGCGACGACGACCCCTTCGCAAGCA 1139
Qy 597 rLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPr 617
Db 1140 GCCCGCTTCTGGGCTAGCATGATGAGGAGCGACGAGTCCCGCGCGATATGTTCTCTCC 1199
Qy 617 oGluIlePheSerPheGluSerThrThrGlyPheThrLeuTrpGlyMetLeuTyrLysPr 637
Db 1200 AGAGATTTTCATTTTCCACGCGCTCGGATGCGGCTCTAGCGCATGATCTACAAGCC 1259
Qy 637 oHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGI 657
Db 1260 CCACGCTTCGAGCGGAGGAGAACCCACCGCTCTCTTTGTATATGAGGCGCCCA 1319
Qy 657 nValGlnLeuValAsnAsnArgPheIysGlyValIysTyrPheArgLeuAsnThrLeuAl 677
Db 1320 GGTGAGTGTGTGAATTAACCTCTCAAGGATCAAGTACTTGGCGGCTCAACACATGGC 1379
Qy 677 aSerLeuGlyTrpValValValIleAspAsnArgGlySerCysHisArgGlyLeuLy 697
Db 1380 CTCCTCGGCTTACCGCGTGTGTGATTGAGCGGCGGCGCTCTCTGTCAGCGGCGCTTCG 1439
Qy 697 sPheGluGlyAlaPheIysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGI 717
Db 1440 GTTCGAGGCGCCCTGAACCAACCAATGGCGCAGGTGAGATCGAGGACCGAGTGGAGGG 1499
Qy 717 yLeuGlnTrpLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGI 737
Db 1500 CTGCGAGTTCGTGGCGGAGAGTATGGCTTTCATCGACCTGAGCGAGTGGCATCATGG 1559
Qy 737 yTrpSerTyrGlyTrpLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheAr 757
Db 1560 CTGGTCTTACGGGGCTCTCTCTCGCTCAATGGGGCTTAATCCACAGCCCGAGTGTCAA 1619
Qy 757 gValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGI 777
Db 1620 GTGGCCATCGGGGTCGCCCGGTCTGATGATGAGCGGCGGCTCTGATGAGCGGCTACACTGA 1679
Qy 777 uArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValaIaMe 797
Db 1680 GCGCTACATGAGCGCTCTGAGAACCAACGAGCGCTATGAGCGGGTTCGGTGGCCCT 1739
Qy 797 tGlnAlaGluIysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAs 817
Db 1740 GCAGCTGAGAGAGCTGCCCAATGAGCCCAACCGCTTCTTATCTCTCAGCGCTTCTGGA 1799
Qy 817 pGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLy 837
Db 1800 CGAAGACGTGCNACTTTTCCACACAACTCTCTCTCCCAACTGATCCGAGCGGAGAA 1859

```

```

Qy 837 sProTyrAspLeuGlnIleTyr 844
Db 1860 ACCTTACCAGCTCCAGGAATAC 1881

RESULT 11
CR627380 3128 bp mRNA linear HTC 22-SEP-2004
LOCUS Homo sapiens mRNA; cDNA DKF2p762F117 (from clone DKF2p762F117).
DEFINITION CR627380
ACCESSION CR627380
VERSION CR627380.1 GI:50949855
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3128)
AUTHORS Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaiipp,A.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Pobo,G., Han,M. and
Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKF2p762F117) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKF2p762F117
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
Location/Qualifiers
1..3128
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p762F117"
/tissue_type="melanoma (MeWo cell line)"
/clone_lib="762 (synonym: hmel2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="dipeptidylpeptidase 9, N-terminus truncated"

gene
1..3128
/gene="DKF2p762F117"
CDS
<12..1712
/gene="DKF2p762F117"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10477.1"
/db_xref="GI:50949856"
/translation="RKTDVSRPTGTGKNPKIALKLAERFQDSOGKIVSTQKELVOP
FSSLPFKVYIARAGWTRDQKIYAWFLDRPQWLQVLPPALFIPSTENEORLAS
ARAVRNQPVYVYEVTVNINVDIIFYFPOSEGEDELCLFLRANECKTGCHLYKV
TAVLKSGQYDSEPSGEDEPKPIKEIALTSGEWEVLARHSGKIWNNEETKLYVF
QGTQDTPLEHLYVVSVEAGEIVRLTTPGFHSHSCSMQNFDMFVSHSVSTPPCVH
YKLSGDDPDLHKQPRFASMMEAASCPDPYVPEIFPHFRSDVRLYGMVYKPHAL
QPKRGHTFVLVYGGPQVQVIVNNSFKIYKLRNLTLASLYAVVITDGRSGCQGLRF
EGALKNQMGVEIIDEQVEGQFVAEKYGFIDLSRAIHGWSYGGFSLMGLHKQVLF
KVATAGAPVTVMAYDTGYTERYMDVPENNQHQEAGSVLHVKEKLPNEPRLILLHG
FLDENVHFEHTNVLVSLIRAGKPYQIQLYPNERHSIRCPESGSEHYEVTLHLLQLEYL
"

Alignment Scores:
Pred. No.: 2,91e-216 Length: 3128
Score: 1991.50 Matches: 359
Percent Similarity: 78.4% Conservative: 85
Best Local Similarity: 63.4% Mismatches: 121
Query Match: 42.4% Indels: 1
DB: 4 Gaps: 1

```


US-10-825-632-1 (1-882) x CR627380 (1-3128)

QY 308 ArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnProLysValThrPhe 327
DB 12 AGGAAGACGGACTCGTATCGGTACCCAGGACAGGCGCMAGAAATCCCAAGATTGCGCTTG 71
QY 328 LysMetSerGluLeuMetIleAspAlaGluGlyArgIleAspValIleAspLysGlu 347
DB 72 AAATCGCTGAGTTCCAGACTGACAGCCAGGCGCAGATCGTCTCGACCCAGGAGAGAG 131
QY 348 LeuLeuGlnProPheGluLeuLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrp 367
DB 132 CTGTGTCAGCGCTTCACTCGCTGTTCCCGAAGGTGGAGTACATCGCCAGGCGCGGTGG 191
QY 368 ThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGln 387
DB 192 ACCCGGATGGCAAAATACGCGCTGGGCCATGTTCTTGGACCGGCCCCAGCGTGGCTCCAG 251
QY 388 IleValLeuIleSerProGluLeuPheIleProValGluAspValMetGluArgGln 407
DB 252 CTCGTCTCTCCCGCGCGCTGTTTCATCCCGGACACAGAGAAATGAGGAGCGCGCTA 311
QY 408 ArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyrGluGluThr 427
DB 312 GCCTCTGCCAGAGCTGCCCGAGGAATGTCAGCGGTATGTGGTGTACGAGGAGGTCA 371
QY 428 AspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSerHis---GluGlu 446
DB 372 AACGTCTGGATCAATGATGATGACATCTTCTATCCCTTCCCAATACAGAGGAGAGAC 431
QY 447 GluLeuGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeuTyrIle 466
DB 432 GAGCTCTGCTTCTCCCGCGCAATGAATGACAGCGGCTTCTGCCATTTGTACAAAGTC 491
QY 467 ThrSerIleLysGluSerLysTyrLysArgSerSerGlyLeuProAlaProSer 486
DB 492 ACCCGCTTTAAATCCCGAGGCTACGATTGGAGTGAGCCCTTCAGCGCCCGGGAAGAT 551
QY 487 AspPheLysCysProLysGluGluIleAlaIleThrSerGlyGluTrpGluValLeu 506
DB 552 GAATTTAAGTGCCTTAAAGGAAGAGATTGCTCTGACCAGCGGTGAATGGAGGTTT 611
QY 507 GlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyrPheGluGly 526
DB 612 GCGAGGACACGGCTCCAAGATCTGGGTCAATGAGGAGACCAAGCTGGTGTACTTCCAGGCG 671
QY 527 ThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsnProGlyGlu 546
DB 672 ACCAAGACACGCGCTGGAGCACCACTCTAGTGGTACGCTATGAGCGCGCGCGAG 731
QY 547 ValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGlnHisCysAsp 566
DB 732 ATCTAGCGCTCACACGCGCGCTTCTCCATAGCTGTCTCANTGAGCCAGAACTTCGAC 791
QY 567 PhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrIleLeu 586
DB 792 ATGTCTGTACGACACTACAGCGCTGACAGCGCGCTGCTGGTGCAGCTCTACAGCTG 851
QY 587 SerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeuAsp 606
DB 852 ACCGCGCCGACGACGACCCCTGCAAGACGACCGCGCTTCTGGGCTAGCATGATGAG 911
QY 607 SerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSerThr 626
DB 912 GCAGCCAGCTGCGCGCGGATATGTTCTTCAGAGATCTTCATTTCCACACGCGCTCG 971
QY 627 GlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyr 646
DB 972 GATGTGCGGCTCTACGCGCATGATCTCAAGCCGCCAGCGCTTGCAGCCAGGAGAAAGAC 1031
QY 647 ProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLys 666

DB 1032 CCCACCGTCTCTTTGTATATATGAGGCGCCCGACGCTGCTGAATAACTCTTCAAA 1091
QY 667 GlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValValValValIle 686
DB 1092 GGCATCAAGTACTTTCGGGCTCAACACACTGGCCCTCTCCCTGAGCTACCGCTGGTGTGATT 1151
QY 687 AspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMet 706
DB 1152 GACGCGAGGGCTCTCTGACAGAGGGCTTCGGTTCCGAAGGGCCCTGAAAAACCAATG 1211
QY 707 GlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAsp 726
DB 1212 GGCAGAGTGGAGATCGAGGACAGGTGGAGGGCTCGAGTTCTGGCCGAGAAATATGCG 1271
QY 727 PheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyrLeuSerLeu 746
DB 1272 TTCATCGACCTGAGCGGATGCGATCCATGGCTGGTCTCTACGGGGCTTCTCTCGCTC 1331
QY 747 MetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThr 766
DB 1332 ATGGGCTTAATCCACAAGCCCGAGGTGTTCAAGGTGGCCATCGCGGTCCCCGGTCA 1391
QY 767 LeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsn 786
DB 1392 GTCTGATGGCTTACGACACAGGATACACTGAGCGCTACATGGACGTCTCTGAGAACA 1451
QY 787 GluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluPro 806
DB 1452 CAGCAGGCTATGAGCGGGTTCGTGGCCCTCCACGTGGAGAGCTGCCAATGAGGCC 1511
QY 807 AsnArgLeuLeuLeuHisGlyPheLeuAspGlnValHisPheAlaHisThrSer 826
DB 1512 AACCGTTCGTATCTCCACGGCTTCTGGAGAAACGTGCACATTTTCCACACAAC 1571
QY 827 IleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGln 846
DB 1572 TTCCTGCTCTCCCAACTGATCCGAGCAGGAGAAACCTTACAGCTCAGATCTACCCCA 1631
QY 847 GluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHis 866
DB 1632 GAGAGACACAGTATTCCTGCTCCCGAGTGGGCGAGCAGCTATGAAGTCACTGCTGTCAC 1691
QY 867 TyrLeuGlnGluAsnLeu 872
DB 1692 TTTCTACAGGAATACCTC 1709

RESULT 12
DQ053109
LOCUS 2767 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens DPP9 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION DQ053109
VERSION DQ053109.1 GI:66899056
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2767)
AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civallo, D.,
White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL (ex) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 2767)
AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civallo, D.,
White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment. Translation starts at the beginning of
 alignment.

FEATURES

source

1. 2767
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 <1..2767
 /gene="DPPP9"
 /locus_tag="HC13651"

gene

ORIGIN

Alignment Scores: 6.47e-188 Length: 2767
 Pred. No.: 1744.00 Matches: 362
 Score: 1744.00 Conservatives: 90
 Percent Similarity: 53.8%
 Best Local Similarity: 43.1% Mismatches: 386
 Query Match: 37.1% Indels: 2
 DB: 11 Gaps: 2

US-10-825-632-1 (1-882) x DQ053109 (1-2767)

QY 35 PheTyrValGluArgTyrSerTrpSerGlnLeuLysLeuAlaAspThrArgLys 54
 DB 245 TTCAGGTGCAGAGCACTGTGGACGGCTCCGGAGCATCATCCAGCGCGCAAG 304
 QY 55 TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
 DB 305 TACTCGGGCCCTCATTTGTCACAAAGGCCGCCACGACTTCAGTTGTGTCAGAAAGCGGAT 364
 QY 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
 DB 365 GAGTCTGGGGCCCACTCCACCGCTCTACTACTCGGAATGCCATATGGACGGCGAG 424
 QY 95 AsnThrLeuPheTyrSerGluIleProLeuThrIleAsnArgAlaAlaValLeuMetLeu 114
 DB 425 AACTCCCTCTCTACTCTGAGATTCCTCCAGAGGTTCGGAAGAGGCTCTGTCTCTG 484
 QY 115 SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
 DB 485 TCCTGGAGACAGATGCTGGATCAATTCAGGCGCACGCCACCACCATGGGGTCTACTCTGG 544
 QY 135 GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp 154
 DB 545 GAGGAGGAGCTGCTGAGGAGCGGAAACGCTCGGGGTCTTCGGCATCACTCTCTACGAC 604
 QY 155 TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys 174
 DB 605 TTCACACGAGAGTGGCTCTCTCTTCAGCGCAGACGACGCTCTTCACCTGCGCGC 664
 QY 175 AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer 194
 DB 665 GACGCGGCAAGACGGCTTCATGGTGTCCCTATGAACCGGTGGAATCAAGACCCAG 724
 QY 195 CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe 214
 DB 725 TGCTCAGGCGCCCGATGGACCCCAAAATCTGCGCTGCGGACCTCTCTCTCTNNNNN 784
 QY 215 IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgLeuThr 234
 DB 785 NNN 844
 QY 235 TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr 254
 DB 845 NNN 904
 QY 255 PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGlu 274
 DB 905 NNN 964

QY 275 ThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer 293
 DB 965 NNNNNNGTTTCAGAGGGCCCTCAAGACGCTGCGAATCCTGTATGAGAGAGTCGATGATCC 1024
 QY 294 GluValGluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPhe 313
 DB 1025 GAGGTGGAGGTTCATTCACGTCCTCTCTGCGCTAGAGAAGAAAGAGACGACTCGTAT 1084
 QY 314 ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMet 333
 DB 1085 CGGTACCCAGGACAGGACAGCAAGATTCACCAAGATTGCTTGAATTCGCTGAGTTCACG 1144
 QY 334 IleAspAlaGluGlyArgIleIleAspValIleAspLysIleLeuIleGlnProPheGlu 353
 DB 1145 ACTGACAGCCAGGCAAGATCGTCTCGACCCAGAGAGAGGTGTGTGAGCCCTTCAGC 1204
 QY 354 IleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyr 373
 DB 1205 TCGTGTTCCTCCAGAGGTGGAGTACATCGCCAGGGCCGGGTGGACCCGGGATGCAATAC 1264
 QY 374 AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerPro 393
 DB 1265 GCCTGGGCATGTTCTTGGACCGGCCCCAGCAGTGGCTCCAGCTCGTCTCTCTCCCGCG 1324
 QY 394 GluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSerVal 413
 DB 1325 GCCCTGTTTCATCCGAGCACAGAGATGAGGAGCAGCGGTAGCCTCTCCAGAGCTGTC 1384
 QY 414 ProAspSerValThrProLeuIleIleTyrGluThrThrAspIleTrpIleAsnIle 433
 DB 1385 CCAGGATGTCAGCCGTATGTGGTGTACGAGGAGGTCAACACCTCTGGATCAATGTT 1444
 QY 434 HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
 DB 1445 CATGACATCTTCTATCCCTTCCCAATCAGAGGAGAGGAGCAGCTCTGCTTCTCCCG 1504
 QY 453 AlaSerGluCysLysThrGlyPheArgHisIleTyrLysIleThrSerIleLeuLysGlu 472
 DB 1505 GCCAATGAATCAAGACCGGCTTCTGCCATTTGTACAAAGTCACCCCGCTTTTAAATCC 1564
 QY 473 SerLysTyrLysArgSerSerGlyLeuProAlaProSerAspPheLysCysProIle 492
 DB 1565 CAGGGCTACGATTGAGTGAGCTTCCAGCCCGGGGAGAGATGAATTAAGTCCCAAT 1624
 QY 493 LysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn 512
 DB 1625 AAGGAGAGATTGCTCTGACCGCGGTGAATGGAGGTTTTGGCGAGGACGCGCTCCAAG 1684
 QY 513 IleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeu 532
 DB 1685 ATCTGGGTCAATGAGGAGACCAAGCTGTGTACTTCCAGGGGCCAAGGACGCGCGCTG 1744
 QY 533 GluHisIleLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp 552
 DB 1745 GAGCACCACTCTAGTGTGAGTATGAGCGCGCGCGGAGATCGTACGCTCACCACG 1804
 QY 553 ArgGlyTyrSerHisSerCysIleSerGlnHisCysAspPhePheIleSerLysTyr 572
 DB 1805 CCGCGCTTCTCCATAGCTGTCCATGAGCCAGNNNNNNNNNNNNNNNNNNNNNNNNNN 1864
 QY 573 SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp 592
 DB 1865 NNN 1924
 QY 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro 612
 DB 1925 NNN 1984
 QY 613 AspTyrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
 DB 1985 NNN 2044
 QY 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIle 652

QY 804 rGluProAsnArgLeuLeuLeuLeuHi sGlyPheLeuAspGluAsnValHisPheAlaHi 824
 DB 759 TGAACCAAACTGTTACTGCTCTTACATGTTCTCTGGATGAGATGCCATTTTGACACA 818
 QY 824 sThrSerLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTy 844
 DB 819 TACAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTACAGATCTA 878
 QY 844 rProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLe 864
 DB 879 TCCTCAGAGACACAGCATAAGAGTTCTGAATCGGAGAACATTATGAATGCACTCT 938
 QY 864 uLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882
 DB 939 TTTGCACCTACCTTCAAGAAAACCTTGATCAGTATTGCTGCTCTAAAGTGATA 993

RESULT 14
 AFI75225 1292 bp mRNA linear HTC 01-AUG-2003
 LOCUS Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.
 DEFINITION
 ACCESSION AFI75225
 VERSION AFI75225.1 GI:33338055
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 1292)
 Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B.,
 Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J.,
 Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zheng, M.Y., Teng, C.Y.,
 Liu, Q., Yu, L.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T.
 Direct Submission
 Submitted (04-AUG-1999) Molecular Medicine Center for
 Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
 Bei Li Shi Lu, Beijing, 100037, P.R. China

FEATURES
 source Location/Qualifiers
 1..1292
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="aorta"
 142..1023
 /codon_start=1
 /product="MSTP135"
 /protein_id="AAQ13650.1"
 /db_xref="GI:33338056"
 /translation="MTOLAKQRNFPFPFMIQQVFLFTILLQKFSLLKVLVLDLHCWGCS
 TSPHGSGLENNILLXLFYGGPKVOLVNNRFGKVFRLNLPFLGVVAVXGIDNRGS
 CPGRLKEGAFKPMGOIEIDDOVGLOYLASRYDFIDLDVRGHWGYSYGLSIMAL
 MQRSDI FRVALAGAPVTLNIFPYGTGYERYMGHPDQNEQGYLGSVMAQAEKFESEPN
 RLHLHGLFDENVHFAVTSILLSFLVRAGKPYDLQIYPOERHSIRVPESGEHYELHL
 HYLOENLGSRIAALKVT"

CDS
 9..31e-186 Length: 1292
 1720..50 Matches: 325
 Percent Similarity: 97.1% Conservative: 5
 Best Local Similarity: 95.6% Mismatches: 8
 Query Match: 36.6% Indels: 2
 DB: 4 Gaps: 1

ORIGIN
 Alignment Scores:
 Pred. No.: 9.31e-186 Length: 1292
 Score: 1720.50 Matches: 325
 Percent Similarity: 97.1% Conservative: 5
 Best Local Similarity: 95.6% Mismatches: 8
 Query Match: 36.6% Indels: 2
 DB: 4 Gaps: 1

US-10-825-632-1 (1-882) x AFI75225 (1-1292)

QY 545 GlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGlnHis 564
 DB 3 GGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAATTTCTGCTGATCAGTCAGCAC 62
 QY 565 CysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyr 584

Db 63 TGTGACTTCTTTATAGTAAGTATAGTAACCAAGAGATCCACACTGTGTGCTTCTTAC 122
 QY LysLeuSerProGluAspAspProThrCysLysThrLysGluPheThrAlaThrIle 604
 DB 123 AAGCTATCAGTCTCTGAAGATGACCCCACTTGCAAAACAAGGAATTTTGGGCCACCAAT 182
 QY 605 LeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSer 624
 DB 183 TTGGATTTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGT 242
 QY 625 ThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu---GlnProGly 643
 DB 243 ACTACTGCAATTTACATTTGTATGGGATGCTCTACAAGCCCTCATGGAGCTACAGGCTGGA 302
 QY 644 LysLysTyrProThr--ValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAs 663
 DB 303 AAGAAATATCTACTGNGGCTGTTTCATATATGGGGGCCCTTAAGGTGCAGTTGTGTGANA 362
 QY 663 nArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVa 683
 DB 363 TCGGTTTAAAGAGATCAAGTTTTTCCCTTGATACCTTACCTCTTTTAGTTATGTGT 422
 QY 683 lValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
 DB 423 TGNAGGGATAGACACACAGGGGATCTGTCCCCGAGGGCTTAAATTTGAAGGCGCTTTAA 482
 QY 703 sTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSe 723
 DB 483 ATTTAAATGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTC 542
 QY 723 rArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrTrpSerTyrGlyGlyTy 743
 DB 543 TCGATATGATTTTACTTACCTTAGATCTGTGGGCATCCACGCTGCTCTCTATGGAGATA 602
 QY 743 rLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAl 763
 DB 603 CCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTGCTGGGGC 662
 QY 763 aProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPr 783
 DB 663 CCCAGTCACTCTGTGGATCTCTTATGATACAGGATACCGGAACGTTATATGGTCAACC 722
 QY 783 oAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePr 803
 DB 723 TGACCAAGATGACACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCC 782
 QY 803 oSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAl 823
 DB 783 CTCTGACCAATCGTTTACTGCTCTTACATGTTCTTCTGATGAGAAATGTCATTTTTCG 842
 QY 823 aHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnI 843
 DB 843 ACATACCAAGTATTTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGAT 902
 QY 843 eTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHi 863
 DB 903 CTATCTCAGGAGAGACACACAGAGTTCTTGAATCGGAGAACATTATGAATGCACTGCA 962
 QY 863 sLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882
 DB 963 TCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTGCTGCTCTTAAAGTGATA 1020

RESULT 15
 BM557438
 LOCUS BM557438
 DEFINITION AGENCOURT_6578992 NIH_MGC_41 Homo sapiens cdna clone IMAGE:5466759
 5', mRNA sequence.
 ACCESSION BM557438
 VERSION BM557438.1 GI:18799430
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 1042)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1CM1967 row: k column: 16
High quality sequence stop: 697.

FEATURES
source

Location/Qualifiers
1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones IMAGE:3466759
/tissue_types="amelanotic melanoma, cell line"
/lab_hosts="NIH_MGC_41"
/note="Organ: skin; Vector: pOT7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 8.27e-183 Length: 1042
Score: 1693.50 Matches: 324
Percent Similarity: 94.0% Conservative: 3
Best Local Similarity: 93.1% Mismatches: 15
Query Match: 36.0% Indels: 6
DB: 3 Gaps: 1

US-10-825-632-1 (1-882) x BM557438 (1-1042)

QY	504	GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr	523
Db	1	GAAGTTCTTGCCCGCATGGATCTATATCCAGTTGATGAAGTCAGAGGCTGGTATAT	60
QY	524	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	543
Db	61	TTTGAGGGCACAAGAGCTCCCTTTAGAGCATCCTGTAGTACGTAGTCAGTTACGTAAT	120
QY	544	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln	563
Db	121	CCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCATCTTGTGTCATCAGTCAG	180
QY	564	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	583
Db	181	CACGTGTGACTCTTTATTAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCCCTT	240
QY	584	TyrLysLeuSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr	603
Db	241	TACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGAAATTTGGGCCACC	300
QY	604	IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu	623
Db	301	ATTTTGGATTACAGAGGTCTCTTCTTGACTACTACTCTCCAGAAATTTTCTTTTGA	360
QY	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643
Db	361	AGTACTACTGGATTACATTGTATGGATGCTCTACAGCCCTCATGATCTACAGCTGGA	420

QY	644	LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAen	663
Db	421	AAGAAATATCTACTGTGCTGCTTCATATATGGTGGTCTCAGGTGAGTGGTGAATAT	480
QY	664	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal	683
Db	481	CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGTT	540
QY	684	ValValIleAspAenArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	703
Db	541	GTAGTGATAGACACACAGGGATCTGTCTACCGAGGGCTTAAATTTGAAGGCCCTTTANA	600
QY	704	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	723
Db	601	TATAAATGGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT	660
QY	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyrGlyTyr	743
Db	661	CGATATGATTTTCAATTGACTTGTAGATCGTGTGGGCATCCAGGCTGGTCTTATGAGGATAC	720
QY	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763
Db	721	CTCTCCCTGATGGCATTAATGCAGAGGTCAAGATATCTTCAGGGTTGCTATTCTGCGGCC	780
QY	764	ProValThrLeuTyrPheThrAspThrGlyTyrThrGluArgTyrMetGlyHisPro	783
Db	781	CAAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGTACGTTATATGGTCACTT	840
QY	784	AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro	803
Db	841	GACCAAGATGAACAGGGCTATTACTTANGATCTGTGGCCATGCAAGCAGAAAGTCCCCC	900
QY	804	SerGluProAsnArgLeuLeuLeuHisGly-PheLeuAspGluAsnVal-HisPheA	823
Db	901	TCTGAACCAATCGATTACTGCTTACATGCTTTTCTGGATGAGAATGTCCCATTTTG	960
QY	823	IaHisThrSer-IleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeu---	841
Db	961	CACATACCCAGGATATTAATCTGGAGTTTNTTGGAGGGGCTGAAAAAGCCCATATGGAT	1020
QY	842	-----GlnIleTyrPro 845	
Db	1021	TTTACACGAATCTATCCT 1038	

Search completed: April 15, 2006, 05:59:28
Job time : 8374 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 00:36:45 ; Search time 379 Seconds
(without alignments)
4136.700 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAATEPEQLGVIFETADC.....HLLHYLOENLGSRAALKVI 882

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlip
-Q-/abss/ABSSWEB spoil/US10825632/runat_14042006_092603_4744/app_query.fasta_1
-DB=Issued Patents NA -QWFAST=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs05p -USER=US10825632 @CGN_1_1_290 @runat_14042006_092603_4744
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	4700	100.0	2671	3	US-09-976-674-2
2	4700	100.0	3120	3	US-10-070-464-2
3	4680	99.6	4829	3	US-09-976-674-12
4	4443.5	94.5	2797	3	US-09-976-594-1103
5	4385.5	93.3	4685	3	US-09-976-674-22
6	4385	93.1	4676	3	US-09-976-674-20
7	4092.5	87.1	4523	3	US-09-976-674-8
8	3661.5	77.9	4309	3	US-09-976-674-14
9	2870	61.1	2617	3	US-09-976-674-4

10	2870	61.1	4219	3	US-09-976-674-28	Sequence 28, Appl	
11	2870	61.1	4302	3	US-09-976-674-24	Sequence 24, Appl	
12	2820.5	60.0	4180	3	US-09-976-674-36	Sequence 36, Appl	
13	2820.5	60.0	4263	3	US-09-976-674-34	Sequence 34, Appl	
14	2649	56.4	4076	3	US-09-976-674-32	Sequence 32, Appl	
15	2649	56.4	4159	3	US-09-976-674-30	Sequence 30, Appl	
16	2599.5	55.3	4037	3	US-09-976-674-40	Sequence 40, Appl	
17	2599.5	55.3	4120	3	US-09-976-674-38	Sequence 38, Appl	
18	2422	51.5	1669	3	US-10-070-464-6	Sequence 6, Appl	
19	1836.5	39.1	1083	3	US-10-070-464-8	Sequence 8, Appl	
20	1645.5	35.0	1197	3	US-09-976-674-26	Sequence 26, Appl	
21	1391	29.6	2411	3	US-09-976-674-10	Sequence 10, Appl	
22	1278	27.2	1356	3	US-09-976-674-18	Sequence 18, Appl	
23	1026.5	21.8	832	3	US-09-976-674-16	Sequence 16, Appl	
24	755.5	16.1	823	3	US-09-280-116-171	Sequence 171, Appl	
C	25	719.5	15.3	14555	3	US-09-902-540-1096	Sequence 1096, Ap
26	719	15.3	2946	3	US-09-902-540-9675	Sequence 9675, Ap	
27	597.5	12.7	4280	3	US-09-079-592-1	Sequence 1, Appl	
28	597.5	12.7	5496	3	US-09-462-284-1	Sequence 1, Appl	
29	572	12.2	2199	3	US-09-518-550-38	Sequence 38, Appl	
30	529	11.3	3407	3	US-10-002-593-5	Sequence 5, Appl	
31	529	11.3	3407	3	US-09-949-016-275	Sequence 5, Appl	
32	529	11.3	3407	3	US-10-423-714-5	Sequence 5, Appl	
33	525	11.2	2924	3	US-09-023-655-1026	Sequence 1026, Ap	
34	525	11.2	2924	6	PCT-US93-07923-1	Sequence 1, Appl	
35	525	11.2	3419	3	US-09-949-016-4579	Sequence 4579, Ap	
36	500.5	10.6	20740	3	US-09-902-540-1223	Sequence 1223, Ap	
37	498.5	10.6	2115	3	US-09-902-540-4518	Sequence 4518, Ap	
38	488.5	10.4	2214	3	US-09-902-540-4124	Sequence 4124, Ap	
39	488.5	10.4	23091	3	US-09-902-540-1204	Sequence 1204, Ap	
C	40	464	9.9	620	3	US-09-976-674-16	Sequence 16, Appl
41	461	9.8	2576	3	US-10-363-937-21	Sequence 21, Appl	
42	461	9.8	2583	3	US-09-976-674-6	Sequence 6, Appl	
43	461	9.8	4541	3	US-09-976-674-42	Sequence 42, Appl	
44	454	9.7	2815	2	US-08-230-491A-1	Sequence 1, Appl	
45	454	9.7	2815	2	US-08-619-280A-1	Sequence 1, Appl	

ALIGNMENTS

RESULT 1
US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akimsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-2

Alignment Scores:
Pred. No.: 0
Score: 4700.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 3
Length: 2671
Matches: 882
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-825-632-1 (1-882) x US-09-976-674-2 (1-2671)

QY	1	MetAlaAlaMetGluThrGluInLeuGlyValGluIlePheGluThrAlaAspCys	20	Db	1088	TATATTGCCAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTGGTCCATCTCTACTAGAT	1147
Db	8	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTGGAAACTGGGACTGT	67	QY	381	ArgSerGlnThrArgLeuGlnIleValIleuSerProGluLeuPheIleProValGlu	400
QY	21	GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTy-rValGluArgTyr	40	Db	1148	CGCTCCAGACTCGCCTGCAGATAGTGTGATCTCAGCTGAATATTATTATCCAGTAGAA	1207
Db	68	GAGGAGAAATATTGAATCACAGGATCGGCCTAAATTGGAGCCTTTTATGTGGCGGTAT	127	QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyRHISGlyTyrMetMet	60	Db	1208	GATGATGTTATGGAAAGGCAGAGACTCATGTGAGTCAGTGCCTGATCTGTGAGCCACATA	1267
Db	128	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAAAATATCATGGCTACATGATG	187	QY	421	IleIleTyrGluGluThrThrAspIleTrpIleAenIleHisAspIlePheHisValPhe	440
QY	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80	Db	1268	ATTATCTATGAAGAAACACACAGATCTGGATAAATATCCATGACATCTTTCATGTGTTTT	1327
Db	188	GCTAAGGACCACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACTCATTTCA	247	QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100	Db	1328	CCCCAAAGTCACGAAGAGGAAATTGAGTATTATTTTGGCTCTGAATGCCAAAAACAGGTTTC	1387
Db	248	GACAGAACTCTATTACCTTGGCATGCTCTGGTGAGAACAGAGAAAAATACACTGTTTATTCT	307	QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120	Db	1388	CGTCAATTTATACAAAATTTACATCTATTATTAAGGAAGCAAAATATAAAGCATCCAGTGGT	1447
Db	308	GAATTTCCAAAAATATCAATAGACGACAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG	367	QY	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140	Db	1448	GGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAGT	1507
Db	368	GATCTTTTTCAGGCAACACATGACTATGGAATGTATTCTCGAGAAGAAACATTTAAGA	427	QY	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	Db	1508	GGTGAATGGGAAGTCTTGGCCGCATGATCTAATATCAAGTTGATGAAGTCAGAAGG	1567
Db	428	GNAGAAAACCATTTGGNACNGTCGNAATTTGCTTTACGATTAATCACCCAGGAAGTGG	487	QY	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	Db	1568	CTGGTATATTTTGAAGCACCACCAAGATCTCCCTTTAGAGCATCACCTGTACGTAGTCA	1627
Db	488	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGGCCACAAGA	547	QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	Db	1628	TACGTAATCTCTGGAGAGGTGACAAAGCTGACAGCCGTGGCTACTCACATTTCTGCTGC	1687
Db	548	TTTAGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG	607	QY	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	Db	1688	ATCAGTCAGCATCTGTGACTTCTTTTAAAGTAAGTATATAGTAACCAAGAAATCCACACTGT	1747
Db	608	GATCCAAAATTTATGCCCTGTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT	667	QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	Db	1748	GTGTCCCTTTTACAAGCTATCAAGTCTTGAAGATGACCCAACTTGCAAAAACAAAGGAATTT	1807
Db	668	TGGATATCTAACATCGTAACCCAGAGAAGAAAGAGACTCACTTATGTGCACAATGAGCTA	727	QY	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	Db	1808	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATACTCTCCAGAAATTTTC	1867
Db	728	GCCACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTTCTCCCAAGAAGA	787	QY	621	SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	Db	1868	TCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTA	1927
Db	788	TTTGATAGATATTCTGGCTATTGGTGTGTCCAAAGCTGNAACACTCCCAAGTGGTGGT	847	QY	641	GlnProGlyLysLysTyrProThrValIleuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Db	1928	CAGCTTGGAAAGAAATATCTCTGCTGTTTCTATATATGTTGGTCTCTCAGTGCAGTTG	1987
Db	848	AAAATCTTTAGAAATCTATATGAAGAAAAATGATGAATCTGAGTGGAAATATTATCATGTT	907	QY	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	1988	GTAATAATATCCGATTTAAAGGAGTCAAGTATTTCGCTTGAATACCTTAGCTCTCTAGGT	2047
Db	908	ACATCCCCATATGTTGGAAACAAGAGGGCAGATTCATTCGTTATCTCTAAAAACAGGTACA	967	QY	681	TyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly	700
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	2048	TATGTGGTGTAGTATAGAACACAGGGGATCTCTGTCCCGAGGGCTTAAATTTGAAGGC	2107
Db	968	GCAATCTCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATTGATGCTGGAAGAGGATC	1027	QY	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Db	2108	GCCTTTAAATATAAAATGGGTCAAAATAGAAATTGACCATCAGGTGGAAGACTCCCAATAT	2167
Db	1028	ATAGATGTCATAGATAAGGAACATAATTCAAACCTTTTGGAGATCTATTGGAAGGAGTTGAA	1087	QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380				

Db 2168 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTAT 2227
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2228 GGAGGATACCTCTCCCTGATGGCATTAATGCGAGGTCTGATATCTTCAGGGTTCCTATT 2287
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2288 GCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACCGAAGCTTATATG 2347
Qy 781 GlyHisProAspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2348 GGTCACTCTACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2407
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2408 AAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGNATGTC 2467
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2468 CATTTTGCACATACCACTATATTACTGAGTTTCTTACTGAGGGCTGGAAAGCCATATGAT 2527
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2528 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTAT 2587
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2588 GAACCTGCATCTTTTGGCACTACCTTCAAGAAAACTTGGATCAGATTGCTGCTCTAATA 2647
Qy 881 Vallie 882
Db 2648 GTGATA 2653

RESULT 2

US-10-070-464-2
; Sequence 2, Application US/10070464
; Patent No.: 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-2

Alignment Scores:

Pred. No.: 0 Length: 3120
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-1 (1-882) x US-10-070-464-2 (1-3120)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGAAACTGCGGACTGT 273

Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATATTGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCCTGGAGTCAGCTTTAAAAGCTGCTTGCAGATACCAAGAAATATCATGGCTACATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGACCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAAATCTATTACCTTGGCCATGCTCTGGTGAAGACAGAGAAATACATCTGTTTATTCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCAAAACACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAAAGCCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAAACCATTTGGAAACAGTCGGAATTCCTTTACGATTATCACCAAGAAAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGAGGGGCCACAAGGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTTACGCAACACCTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCACATACCGGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGGATTCCTTTATACATAGACAGCATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAACATCTGTAAACAGAGAGAAAGGAGACTCACTTATGTGCACATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTCTGGCTATTGGGTGCTCCAAAAGCTGAAACAACTCCAGTGGTGGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 1054 AAAATTTCTTAGAAATTTCTATGCAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCTTATGTTGGAACAAGAGGCGCAGATTCAATTCCTGTTATCTCTTAAACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAAAATCTTAAAGTCATCTTTTAAAGATGTGAGAAATATGATGTGATGTGAGGAGGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCATAGTAAGGAACATAATTCACCTTTTGAGATTCCTATTGAGGAGTTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGACTGGATGGACTCTGAGGGGAAATATGCTGGTCCATCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400

Db	1354	CGCTCCAGACTCGCCTACAGATAGTGTGGATCTCACTGAAATATTTATCCAGTAGAA	1413	Db	2434	GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTCAGATATCTTCAGGGTTGCTATT	2493
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Qy	761	AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	1414	GATGATGTTATGGAAAGCGCAGAGACTCATTGAGTCAGTGCCTGATTCGTGACGCCACTA	1473	Db	2494	GCTGGGGCCCACTCTCTGTGGATCTTTATGATACAGGATACACGGACCTATATATG	2553
Qy	421	IleIleTyrGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe	440	Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	1474	ATTATCTATGAAGAAACAACAGACATCTGGATTAATATCATGACATCTTTCACTGTTTT	1533	Db	2554	GGTCAACCTTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2613
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysIysThrGlyPhe	460	Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	1534	CCCCAAGTACGAAGAGAAATGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC	1593	Db	2614	AAGTTCCCTCTGAACCAATCGTTTACTCTCTTACATGGTTTCCTGGATGAGAATGC	2673
Qy	461	ArgHisLeuTyrIysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	1594	CGTCATTTATCAAAATTTACATCTATTATTAAGGAAGCAAAATATAAAGCATCCAGTGGT	1653	Db	2674	CATTTTGACATACACAGTATATTACTAGTGTTTTGTAGTGGGCTGGAAGCCCATATGAT	2733
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	1654	GGGCTGCCTGCTCCAAAGTGATTTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTACCAGT	1713	Db	2734	TTACAGATCTATCTCTCAGGAGACACAGACATAAGAGTTCTCCTGAATCGGAGAACATTAT	2793
Qy	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	1714	GCTGAATGGGAAGTCTCTGGCGGCATGGATCTAAATATCAAGTTTGATGAAGTCAGAAAG	1773	Db	2794	GAACTGCATCTTTTGCACTACCTTCAAGAAAACTTGGATCACGTATTGCTGCTCTAAAA	2853
Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540	Qy	881	ValIle 882	
Db	1774	CTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTGCTAGTCACT	1833	Db	2854	GTGATA 2859	
Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	RESULT 3			
Db	1834	TACGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACATCTTGTCTGC	1893	US-09-976-674-12			
Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	; Sequence 12, Application US/09976674			
Db	1894	ATCAGTCAGCACTGTGACTCTTTTATAAGTAAAGTATAGTAACACAGAAATCCACACTGT	1953	; Patent No. 6844180			
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysIysThrIysGluPhe	600	; GENERAL INFORMATION:			
Db	1954	GTGTCCTTTACAGCTATCAGTCTCGAAGATGACCACTGCAAAACAAAGAAATTT	2013	; APPLICANT: OJ. Steve			
Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	; APPLICANT: Akinsanya, Karen			
Db	2014	TGGGCCCACTTTTGGATTTCAGCAGGTCCTTCTCTGACTATACTCTCCAGAAATTTTC	2073	; APPLICANT: Riviere, Pierre			
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	; APPLICANT: Jumein, Jean-Louis			
Db	2074	TCCTTTGAAAGTACTACTGGATTATACATTGTATGGGATGCTCTCAAGACCTCATGATCTA	2133	; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
Qy	641	GlnProGlyLysIysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660	; FILE REFERENCE: 70669			
Db	2134	CAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCAGGTGCAGTTG	2193	; CURRENT FILING DATE: 2001-10-12			
Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680	; PRIOR APPLICATION NUMBER: US/09/976,674			
Db	2194	GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTTAGGT	2253	; PRIOR FILING DATE: 2000-10-17			
Qy	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700	; NUMBER OF SEQ ID NOS: 61			
Db	2254	TATGTGTTGTAGTATAGACACACAGGGATCTGTGTCACCGAGGCTTAAATTTGAAGGC	2313	; SOFTWARE: Patentin version 3.1			
Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGluLeuGluTyr	720	; SEQ ID NO 12			
Db	2314	GCCTTTTAATATAAATGGTCAATAGAAATGACGATCAGTTCGAAGGACTCCATAT	2373	; LENGTH: 4829			
Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740	; TYPE: DNA			
Db	2374	CTAGCTCTCGATATGATTTCACTTGTAGTCTGTGGGATCCAGGCTGCTCTAT	2433	; ORGANISM: Homo sapiens			
Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760	US-09-976-674-12			
				Alignment Scores:			
				Pred. No.:	0	Length:	4829
				Score:	4680.00	Matches:	882
				Percent Similarity:	99.8%	Conservative:	0
				Best Local Similarity:	99.8%	Mismatches:	0
				Query Match:	99.6%	Indels:	2
				DB:	3	Gaps:	0
				US-10-825-632-1 (1-882) x US-09-976-674-12 (1-4829)			
Qy				1	MetaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20	
Db				214	ATGGCAGCAGCAATGGAAACAGAACAGCTGGTGGTGGATATTGAAACTCGGACTGT	273	
Qy				21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40	
Db				274	GAGGAGATATTGATACAGGATCGGCTAAATTTGAGCCCTTTTATGTTGAGCGGTAT	333	
Qy				41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60	

Db 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACAGAAATATCATGCTACATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
Qy 81 AspArgLysLeuTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAATCTATTACCTTGGCCATGCTGGTGAGAACAGAGAAATACATGTTTATTCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGGAGGCTCTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAAGAAGAACTATTAA 633
Qy 141 GluArgLysArgLysGlyThrValGlyLysSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAACCGATTTGGAACTGCGAATTCCTTTACGATTTATCACCAGGAAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAGAATGGAGGGCCACAG 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAATATTATGCTGCTGCTGATCCAGACTGGATGCTTTTATATACATAGCAAC 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAACATCGTNACAGAGAAGAAAGAGACTCATTTATGTGCACATAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 934 GCCAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGCTCCCAAGAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTTGGCTATTGGTGTGCTCCAAAGCTGAAACAACTCCCACTGGTGT 1053
Qy 281 LysIleLeuArgLysLeuTyrGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 1054 AAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCTATGTTGGAACCAAGAGGGCAGATTCATTCGTTATCTTAAACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAATCTCTAAAGTCACCTTTAAGATGTCAGAAATAATGATTCATGCTCGAAGAGGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCATAGATAAGGAATAATCAACCTTTTGGAGATTCATTTGAAGGAGTTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGACTGGATGACCTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAGGACAGACTCAATTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473

Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1534 CCCCAAGTCACAGAGAGAAATGAGTTATTTTTCCTCTGATGCCAACAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGCTCGCTCCCAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAGT 1713
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTTCTTGGCGGATGGATCTAATATCCAGTTGATGAGTCAGAGG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TAGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGC 1893
Qy 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCATCTGTGACTTCTTTAATAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCCTTTACAAGCTATCAAGTCTCGAAGATGACCCCACTTGCAGAAACAAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
Db 2014 TGGGCGACCATTTTGGATTTCAGCAGGTCTCTCTGCTGACTATATCTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL 660
Db 2134 CAGCCTGGAAAGAAATATCTTACTGTGCTGTTCAATATATGGTGTCTCTCAGTGCAGT 2193
Qy 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
Db 2194 TGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAG 2253
Qy 680 LysTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
Db 2254 GTTATGTGGTGTGTAGTATAGCAACAGGGGATCCTGTCCCGAGGGCTTTAAATTTGAAG 2313
Qy 700 LysAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln 720
Db 2314 GGGCTTTAATATATAAATGGTCAATAGAAATTTGACGATCAGTGGGAAGGACTCCAAT 2373
Qy 720 TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSer 740
Db 2374 ATCTAGTCTCTCGATATGATTTTCTTACTTACTGTAGATCTGTGGGCATCCACGGCTGTCT 2433
Qy 740 TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 2434 ATGGAGGATACCTCTCCCTGATGCAATTAATGCAAGAGGTGAGATATCTTCAGGGTTGCTA 2493
Qy 760 LeuAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrM 780
Db 2494 TTGCTGGGGCCCCAGTCTCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATA 2553

QY	780	etGlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG	800
Db	2554	TGGGTCAACCTCCAGCAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAG	2613
QY	800	lulysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAenV	820
Db	2614	AAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAATG	2673
QY	820	alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA	840
Db	2674	TCCATTTGGACATACCATGATATATCTAGTTTCTAGTGGGCTGGAAGCCATATG	2733
QY	840	spLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisT	860
Db	2734	ATTTACAGATCTATCTCCAGAGACACAGATGATTTCTGAAATCGGGAGACATT	2793
QY	860	YrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuL	880
Db	2794	ATGAACATGCATCTTTTGGCACTACTCTCAAGAAACCTTGGATCAGTATTGCTCTTAA	2853
QY	880	YsValIle 882	
Db	2854	AAGTGATA 2861	
RESULT 4			
US-09-976-594-1103			
; Sequence 1103, Application US/09976594			
; Patent No. 6673549			
; GENERAL INFORMATION:			
; APPLICANT: Furness, Michael			
; APPLICANT: Buchbinder, Jenny			
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS			
; FILE REFERENCE: PA-0041 US			
; CURRENT APPLICATION NUMBER: US/09/976,594			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: 60/240,409			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 1143			
; SOFTWARE: PERL Program			
; SEQ ID NO 1103			
; LENGTH: 2797			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1			
US-09-976-594-1103			
Alignment Scores:			
Pred. No.:	0	Length:	2797
Score:	4443.50	Matches:	840
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	94.5%	Indels:	3
DB:	3	Gaps:	1
US-10-825-632-1 (1-882) x US-09-976-594-1103 (1-2797)			
QY	42	TpSerGlnLeuLysValLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMetAla	61
Db	3	TGGATGACCTTTAAAAGCTGCTTCCGATACCAAGAAATATCATGCTACATGATGGCT	62
QY	62	LysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSerAsp	81
Db	63	AAGGCACCATGATTTTCAATGTTGTGAAGAGATGATCCAGATGACCTCATTCCAGAC	122
QY	82	ArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnArgGluLeuPheTyrSerGlu	101
Db	123	AGAATCTATTACTTGGCCATGCTCGTGAGAACACAGAAATACATCTGTTTATTCTGAA	182
QY	102	IleProLysThrIleAsnArgAlaValLeuMetLeuSerTyrLysProLeuLeuAsp	121
Db	183	ATTCCCAAAACTCATAGACGACGCTCTTAATGCTCTCTTGGAGGCTCTTTTGGAT	242
QY	122	LeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArgGlu	141
Db	243	CTTTTTCAGGCACACTGGACTATGAAATGATTTCTCGAGAAGAAGAACTATTAAAGAA	302
QY	142	ArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGlyThr	161
Db	303	AGAAAACGCATTTGGAACAGTCGGAATTTGCTTTCTACGATATATCACNAGGAAGTGAACA	362
QY	162	PheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGlyLys	181
Db	363	TTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAAAGATGGAGGGCCACAGATTT	422
QY	182	ThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAsp	201
Db	423	ACGCAACAAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGCCCAACATACGGATGGAT	482
QY	202	ProLysLeuCysProAlaAspProAspTyrPheAlaPheIleHisSerAsnAspIleTyrP	221
Db	483	CCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTATATACATAGCAACATATTTGG	542
QY	222	IleSerAsnIleValThrArgGluGluArgGluThrTyrValHisAsnGluLeuAla	241
Db	543	ATACTTAACATCGTAACACAGAGAAGAGAGACTCACTTATGTGCACAATGAGCTAGCC	602
QY	242	AsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPhe	261
Db	603	AAATGGAAGAAGAGATGATGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTT	662
QY	262	AspArgTyrSerGlyTyrTyrTyrCysProLysAlaGluThrThrProSerGlyGlyLys	281
Db	663	GATAGATATCTGGCTATTGGTGTGTCCAAAGCTGAAACAACCTCCAGTGGTGTAATA	722
QY	282	IleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThr	301
Db	723	ATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATCATGTTACA	782
QY	302	SerProMetLeuGluThrArg-ArgAlaAspSerPheArgTyrProLysThrGlyThrAl	321
Db	783	TCCCTATGTTGGAAACAAGCAGGCGAGATTCATTCCTGTTATCCATAAACAGGTACAGC	842
QY	321	aAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleI	341
Db	843	AAATCTTAAGTCACTTTTAAGATGTGAGAAATAAATGATGATGCTGAAGGAAGGATCAT	902
QY	341	eAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyr	361
Db	903	AGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGATTTCTATTTGAAGGAGTTGAATA	962
QY	361	rIleAlaArgAlaGlyTyrThrProGluGlyLys- ---TyrAlaTyrSerIleLeuLeuA	380
Db	963	TATTCAGAGCTGGATGGACTCCTGAGGGAATAAGTAGTGTGCTTGGTCCATCTACTAG	1022
QY	380	spArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValG	400
Db	1023	ATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAG	1082
QY	400	luAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProL	420
Db	1083	AAAGATGATGTTATGGAAGGCGAGACTCATTTGAGTCAGTCCCTGATTTCTGTGACGCCAC	1142
QY	420	eulleIleTyrGluGluThrThrAspIleTyrPheIleAsnIleHisAspIlePheHisValP	440
Db	1143	TAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCTGTTT	1202
QY	440	heProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyP	460
Db	1203	TTCCCAAAAGTCAACGAAGAGAAATTTGATTTTATTTTTGCTCTGAAATGCAAAACAGTT	1262
QY	460	heArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerg	480
Db	1263	TCCGTCTATTATACAAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTG	1322

Qy 480 lYcLYLeuProAlaProSerAspPheLysCysPheLysGluGluLeuAlaLeuThrS 500
Db 1323 GTGGGCTGCCTGCTCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTAACA 1382
Qy 500 eRGLyLutrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgA 520
Db 1383 GTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAA 1442
Qy 520 rGLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValS 540
Db 1443 GGCTGTATATTTGAAGGACCACCAAGACTCCCTTTAGAGCATCCTGTAGTAGTCA 1502
Qy 540 erTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysC 560
Db 1503 GTTACGTAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTACATCTTGTCT 1562
Qy 560 ystleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisC 580
Db 1563 GCATCAGTCAGCACTGTGACTTCCTTTAATAGTAAGTATAGTAACCAAGAGATCCACACT 1622
Qy 580 ySValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluP 600
Db 1623 GTGTGTCCTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTCGCAACCAAGGAAT 1682
Qy 600 heTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIleP 620
Db 1683 TTTGGGCCACCAATTTGGATTTCAGCAGGTCTCTCTCTGACTACTCTCTCCAGAAATTT 1742
Qy 620 heSerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspL 640
Db 1743 TCTCTTTTGAAGTAGTACTACTGGATTTACATTTGATGGGATGCTCTCAAGCCCTCATGATC 1802
Qy 640 euGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnL 660
Db 1803 TACAGCTCGAAAGAAATATCTACTGCTGCTTCATATATGCTGCTCTCAGTGCAGT 1862
Qy 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
Db 1863 TGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAG 1922
Qy 680 lYTrpValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
Db 1923 GTTATGTGTGTGTAGTAGTAGACAACAGGGGATCCTGTCCAGGGGCTTTAAATTTGAAG 1982
Qy 700 lYAlaPheLysTyrLysMetGlyGlnIleAspAspGlnValGluGlyLeuGlnT 720
Db 1983 GCCTCTTTAAATATAAATGGGTCAATAGAAATGACATGAGTGGGAAGGACTCCAT 2042
Qy 720 YrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerT 740
Db 2043 ATCTAGCTCTCGATATGATTTTCATTTAGACTTAGATCGTGTGGGCATCCAGGCTGCTCT 2102
Qy 740 YrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 2103 ATGGAGGATACCTCTCCCTGATGGCATTAATATGACAGGTCCAGATATCTTCAGGGTTCCTA 2162
Qy 760 lealagLYAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrW 780
Db 2163 TTGCTGGGGCCCGACCTCTCTGGATCTCTATATGATACAGGATACACCGAAGCTTATA 2222
Qy 780 etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG 800
Db 2223 TGGGTACCTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGCAG 2282
Qy 800 lulyPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
Db 2283 AAAAGTTCCCTCTCGAACCAATCGTTTACTGCTTACATGTTTCTCTGGATGAGAATG 2342
Qy 820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA 840
Db 2343 TCCATTTTGCACATACCATATATTACTGAGTTTTTTTAGTGAGGGCTGGAAGGCCATATG 2402
Qy 840 spLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisT 860

Db 2403 ATTACAGATCTATCTCAGAGAGACACAGCATPAAGATTCTGATCGGGAGACATT 2462
Qy 860 YrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuL 880
Db 2463 ATGAACCTGCATCTTTTGACCTACCTTCAAGAAAAACCTTGGATCAGGTATTGCTGCTCTAA 2522
Qy 880 ySValIle 882
Db 2523 AAGTGATA 2530
RESULT 5
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-22
Alignment Scores:
Pred. No.: 0 Length: 4685
Score: 4385.50 Matches: 834
Percent Similarity: 94.6% Conservative: 0
Best Local Similarity: 94.6% Mismatches: 1
Query Match: 93.3% Indels: 48
Gaps: 1
US-10-825-632-1 (1-882) x US-09-976-674-22 (1-4685)
Qy 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGCAGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTCGGAGCTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATATTGAATCAAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCTGTGAGTCAGCTTAAAGGCTGCTTCCCATATCCAGAAATAATCATGCTCATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGCACCACCATGATTTTCATGTTGTGAAGGAGATGATCCAGATGACCTCATCA 453
Qy 81 AspArgIleTyrTyrIleAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAAATCTATTACCTTCCCATGCTGTGTCGAGAACAGAGAAAAATACACTGTTTATTCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGCAGCACTCTGAGCTATGGAATGATTTCTCGAGAGAGAGAACTATTAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

Db	634	GAAAGAAACGCGATTCGGAACAGTCGGAATTCGCTTACGATATATCAACCAAGGAAGTGA	693	1714	GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCAGAAGG	1773
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	521	LeuValTyrPheGluCluThyLysAspSerProLeuGluHisHisLeuTyrValValSer	540
Db	694	ACATTTCTGTTTCAAGCCGGTAGTGGATTTATCATCGTAAAAAGATGGAGGCCACAAAGGA	753	1774	CTGGTATATTTTGAAGCACCACAAAGCTCCCTCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
Qy	181	PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet	200	541	TyrValAenProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	754	TTTACGCCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCTCCCAACATACGGATG	813	1834	TAGCTAAATCCTGGAGAGGTGACAGGCTGACTGACCTGGCTACTACATCTCTTGCTGC	1893
Qy	201	AspProLysLeuCysProAlaAspProAenTolIleAlaPheIleHisSerAsnAspIle	220	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCys	580
Db	814	GATCCAAATATATGCTGCTGCTGATCCAGACTCGATCGATTTTATACATAGCAACGATAT	873	1894	ATCAGTCAGACACTGTGACTCTTTATAAGTAAGTATAGTAACCAAGAGATCCACTGT	1953
Qy	221	TrpIleSerAenIleValThrArgGluGluArgArgLeuTyrValHisAenGluLeu	240	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	874	TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCATTATGTGCACATAGACTA	933	1954	GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAGGAATTT	2013
Qy	241	AlaAenMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	934	GCCAACTGGAAGAGATGCCAGATCAGCTGGAGTCGCTTACCTTTGTTCTCCAAGAGAA	993	2014	TGGGCCACCATTTTGGATTTCAGT-----	2036
Qy	261	PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly	280	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	994	TTTGATAGATATTCGCTATTTGGTGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGT	1053	2036	-----	2036
Qy	281	LysIleLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleHisVal	300	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	1054	AAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTATCATGTT	1113	2037	-----CCTCAGGTGCAGTTG	2051
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	661	ValAenAenArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGly	680
Db	1114	ACATCCCTATGTTGAAACCAAGAGGCGCAGATTCATTCGTTATCTCTAAACAGGTACA	1173	2052	GTGAATATCGTTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCCTCTCTAGGT	2111
Qy	321	AlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	681	TyrValValValIleLeuAenAenArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	1174	GCAATCTTAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTGAAGAGAGGATC	1233	2112	TATGTGGTTGTAGTGATAGACACACAGGGGATCTGTCCAGGGGCTTAAATTTGAAGGC	2171
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	1234	ATAGATGTCATAGATAGAGAACTAATTCACACCTTTTGAGATCTATTTGAAGCAGTTGAA	1293	2172	GCCTTTAAATATAAATGGGTCAANTAGAAATTTGACGATCAGTGGAGGACTCCAATAT	2231
Qy	361	TyrIleAlaArgAlaGlyThrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyLysTrpSerTyr	740
Db	1294	TATATTTGCCAGACTGGATGACTCTCTGAGGGAATAATGCTTGGTCCATCTACTAGAT	1353	2232	CTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCACGGCTGGCTAT	2291
Qy	381	ArgSerGlnThrArgLeuGlnIleValIleLysSerProGluLeuPheIleProValGlu	400	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	1354	CGCTCCAGACTCGCTACATAGTGTGTGATCTCAGCTGAATATTATTTATCCAGTAGAA	1413	2292	GGAGGATACCTCTCCCTGATGGCATTAATGGAGAGGTACAGATATCTCAGGGTTGCTATT	2351
Qy	401	AspAspValMetGluArgGlnArgIleGluSerValProAspSerValThrProLeu	420	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	1414	GATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATCTGTGACGCCACTA	1473	2352	GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGATACAGGACGTTATATG	2411
Qy	421	IleIleTyrGluGluThrThrAspIleTrpIleAenIleHisAspIlePheHisValPhe	440	781	GlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	1474	ATTATCTATGAGAAGAACACAGACTCGGATAAATATCCATGACATCTTTCAIGTTTTT	1533	2412	GGTCACCTCGACAGAAATGAACAGGGCTATTACTTAGGTATTTAGTGGGGCTGCAAGCAGAA	2471
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	801	LysPheProSerGluProAenArgLeuLeuLeuLeuHisGlyPheLeuAspGluAenVal	820
Db	1534	CCCCAAAGTCAGGAAGAGAAATTGAGTTTATTTTGCCTCTGAATGCAAAACAGTTTC	1593	2472	ASGTTCCCTCTGAAACCAATCGTTACTGTCTTACATGGTTTCTCTGGATGAGAATGTC	2531
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	1594	CGTCAITTTATACAAATTAACATCTATTTTAAAGGAAAGCAAAATAAACAATCCAGTGT	1653	2532	CATTTTGCACATACACAGTATATTACTAGATTTTACTAGGGGTGGAAGCCATATGAT	2591
Qy	481	GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer	500	841	LeuGlnIleTyrProGluGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	1654	GGCTGCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGT	1713	2592	TTACAGATCTATCTCCAGGAGAGACACAGATTAAGAGTTCTGTAATCGGAGAACATAT	2651
Qy	501	GlyGluTrpGluValLeuGluValArgHisGlySerAenIleGlnValAspGluValArg	520	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaLeuLys	880
				2652	GAACTGCATCTTTTGCNCTACCTTCAAGAAACCTTGGATCACGTATTTGCTGCTCTAAAA	2711


```
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TACGTAAATCTCGAGAGGTGACAGAGCTGACTGACCGTGGCTACTACATCTTGTCTGTC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGCACTGTGACTCTCTTATAAGTAGTAGTAACACAGAAAGAAATCCACACTGT 1953
QY 591 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCCCTTTACAGACTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
DB 2014 TGGGCCACCAATTTGGATTTCAGCAGGTCTCTCTCTGACTATACTCTCCAGAAATTTTC 2073
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2074 TCTTTTGAAGTACTACTGGATTTACATTGTATGGATGCTCTCAACAGCCTCATGATCTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 2134 CAGCTTGGAAAGAAATATCTACTGTCTCTCATATATGTTGGTCTG----- 2180
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB 2180 ----- 2180
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
DB 2180 ----- 2180
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
DB 2181 -----GGTCAAAATAGAAATTTGACGATCAGGTGGAGAGACTCCAAATAT 2222
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
DB 2223 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGTCTAT 2282
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
DB 2283 GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTCGTATT 2342
QY 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
DB 2343 GCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2402
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
DB 2403 GGTCACTCTGACCAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2462
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
DB 2463 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTC 2522
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
DB 2523 CATTTTGCATACACGATATATTAATGAGTTTCTAGTGGGCTGGAAGCCATATGAT 2582
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
DB 2583 TTACAGATCTATCTTCAGAGAGACACAGATAGAGTTCTGTAATCGGAGAACATATAT 2642
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
DB 2643 GAACTGCATCTTTTGCATCTTCAAGAAACCTTTGGATCAGGTATTCGTCTCTCAAA 2702
QY 881 ValIle 882
DB 2703 GTGATA 2708
```

RESULT 7

```
US-09-976-674-8
; Sequence 8, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-8

Alignment Scores:
Pred. No.: 0 Length: 4523
Score: 4092.50 Matches: 780
Percent Similarity: 88.4% Conservative: 0
Best Local Similarity: 88.4% Mismatches: 1
Query Match: 87.1% Indels: 102
DB: 3 Gaps: 1

US-10-825-632-1 (1-882) x US-09-976-674-8 (1-4523)
QY 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 214 ATGGCAGCAGCATATGGAAACACAGACAGCTGGGTGTGGATATTTGAACCTCGGACTGT 273
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 274 GAGGAGAAATATTAATGATCAGATCGGCCTAAATGGAGCCTTTTATGTGTGACCGGTAT 333
QY 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCTTGAGTACGTCTTAAAGAGCTCTTGCCTATCCAGAAATATCATGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCACCACTATGTTTGTGAAGAGGAATGATCCAGATGGACCTCAATCA 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAAATCTATTACCTTGCCTATGCTGTGAGAGAACAGAGAAATACACCTGTTTATTCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
DB 514 GAAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAAAGCCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 574 GATCTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAGAGAAAGAACTATTAAAGA 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAACCGCATTTGGACAGCTCGAATTCCTCTTACGATTATACCAAGAAAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAGCCGGTAGTGGATTTATCCGTAAAGATGAGGGCCCAAGGA 753
QY 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTACGCAACAACTTTTAAAGCCCACTAGTGGAAACTAGTTCTCCCAACATACGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
```

814 Db GATCCAAAATATGCGCTGCTGATCCAGACTGGGATGCTTTATATACATAGCAACGATATT 873
221 Qy TTPILeserAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
874 Db TGGATATCTAAATCGTAACAGAGAGAGAGAGACTCACTTATGTGACAAATGAGCTA 933
241 Qy AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGlu 260
934 Db GCCAAATCGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA 993
261 Qy PheAspArgTyrSerGlyTyrTTPCysProLysValGluThrThrProSerGlyGly 280
994 Db TTTGATPAGATATCTTGGGCTATTGGGTGTCAAAAGCTGAAACAACTCCAGTGGTGT 1053
281 Qy LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
1054 Db AAAAAATCTTAGAATCTATATGAGAAAATGATGATCTGAGGTGGAAATTTATCATGTT 1113
301 Qy ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
1114 Db ACATCCCTTATGTTGMAACAGAGGAGGCGAGATTCATTCGGTTATCTTAAACAGGTACA 1173
321 Qy AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluArgIle 340
1174 Db GCAAAATCTTAAAGTCACATTTAAGATGTGAGAAATAATGATGATGCTGAAGGAAGATC 1233
341 Qy IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
1234 Db ATAGATGTCTATAGATAAGNACTAATCAACCTTTTGAGATCTATTTGAGAGGTGAA 1293
361 Qy TyrIleAlaArgAlaGlyTTPThrProGluGlyLysTyrAlaTTPSerIleLeuLeuAsp 380
1294 Db TATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAATATGTTGGTCCATCTCTACTAGAT 1353
381 Qy ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
1354 Db CGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCTGAAATATTATTATCCAGTAGAA 1413
401 Qy AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
1414 Db GATGATGTATTGGAAGGAGAGAGACTCAATGAGTCAGTGCCTGATTCGTGAGCCCACTA 1473
421 Qy IleIleTyrGluGluThrThrAspIleTTPIleAsnIleHisAspIlePheHisValPhe 440
1474 Db ATTATCTATGAGAAACACACACATCTGGATTAATATCCATGACATCTTTTCATGTTTT 1533
441 Qy ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
1534 Db CCCCAGAGTCAGCAGAGGAAATTCAGTATTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593
461 Qy ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
1594 Db CGTCAITTTATACAAAATTACATCTATTTTAAAGGAAGCAAAATATAACCATCCAGTGGT 1653
481 Qy GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
1654 Db GGCGTCGCTCCTCAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGCAATACCAGT 1713
501 Qy GlyGluTTPGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
1714 Db GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
521 Qy LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
1774 Db CTGGTATATTTTGAAGGCCCAAGAGACTCCCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
541 Qy TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
1834 Db TACGTAAATCCTCGAGAGGTGACAGGCTGACTGACCGGTGCTACTCACATCTTGTCTGC 1893
561 Qy IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580

1894 Db ATCAGTCCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAATCCACACTGT 1953
581 Qy ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
1954 Db GTGTCCCTTTTACAAGCTATCAAGTCTGAGATGCCCAACTTGCAGAAAACAAAGGAATTT 2013
601 Qy TTPAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
2014 Db TGGGCCACCATTTTGGATTTCAGAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
621 Qy SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
2074 Db TCTTTTGAAGTACTACTTGGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA 2133
641 Qy GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
2134 Db CAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGT----- 2178
661 Qy ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
2178 Db ----- 2178
681 Qy TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
2178 Db ----- 2178
701 Qy AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
2178 Db ----- 2178
721 Qy LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
2178 Db ----- 2178
741 Qy GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
2179 Db -----CGGTT-GCTATT 2189
761 Qy AlaGlyAlaProValThrLeuTTPIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
2190 Db GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2249
781 Qy GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
2250 Db GGTCACTCCCTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2309
801 Qy LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
2310 Db AGTTCCTCTCTGAACCAATCGTTTACTTCTTACATGTTTCTTCTGGATGAGATGTC 2369
821 Qy HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
2370 Db CATTTTGCATACATACAGTATATATTACTAGTTTTTTAGTGGGCTGGAAAGCCATATGAT 2429
841 Qy LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
2430 Db TTACAGATCTATCTCTCAGGAGACACACAGATTAAGATTCCTGTAATCGGAGAACATAT 2489
861 Qy GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
2490 Db GAACTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTTATGCTGCTTAAAA 2549
881 Qy ValIle 882
2550 Db GTGATA 2555

RESULT 8

US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen

Tue Apr 18 08:18:33 2006

APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis
FILE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976,674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 4309
TYPE: DNA
ORGANISM: Homo sapiens
US-09-976-674-14

Alignment Scores:
Pred. No.: 0 Length: 4309
Score: 3661.50 Matches: 708
Percent Similarity: 80.1% Conservative: 9
Best Local Similarity: 80.1% Mismatches: 1
Query Match: 77.9% Indels: 176
DB: 3 Gaps: 1

US-10-825-632-1 (1-882) x US-09-976-674-14 (1-4309)

QY	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
DB	214	ATGGCAGCAGCAATGGAAACAGAACACAGCTGGGTGTGGAGATATTTGAAACCTGCGGACTGT	273
QY	21	GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
DB	274	GAGGAGAATATTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTGGAGCGGTAT	333
QY	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
DB	334	TCTCGAGTCAGCTTAAAGAGCTGCTTGGCGATACACAGAAATATCATGGCTACATGATG	393
QY	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
DB	394	GCTAAGGACCCACATGATTTTCATGTTGTGAGAGGAATGATCCAGATGGACCTCATTTCA	453
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
DB	454	GACAGAATCTATTACCTTGGCCATGTCGTGGAGACAGCTTTAATGCTCTTTGGAGGCTCTTTTG	513
QY	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu	120
DB	514	GAAATTCACAAACTATCAATAGACGACGAGCTTAATGCTCTTTGGAGGCTCTTTTG	573
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
DB	574	GATCTTTTTCAGGCACACTGGACTATGGAAATGTTATTCGAGAGAAGAACTATTAGA	633
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
DB	634	GAAAGAAAACGGATTTGGAAACAGTCGGAATTTATCAGCAATATATCAGCAATATCAGCAAT	693
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
DB	694	ACATTTCTCTTTCAAGCGGTAGTGAATTTATCAGCTTAAAGATGGAGGCGGCACAAAGA	753
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
DB	754	TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATG	813
QY	201	AspProLeuLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
DB	814	GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAAGATATT	873
QY	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
DB	874	TGGATATCTATACATCTGTAAACAGAGAAAGAGGAGACTCACTTATGTGCACATGAGCTA	933

QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
DB	934	GCCAAATGGAAGAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCNAGAAGAA	993
QY	261	PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly	280
DB	994	TTTGATAGATATTCTGGCTATTGTGGTGTCCAAAGCTGAAACCACTCCCACTCCGAGTGGT	1053
QY	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
DB	1054	AAAAATCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATCTGTT	1113
QY	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
DB	1114	ACATCCCTATGTTGGAACAAGAGGAGGAGATTCATTCCTTATCCTTAAACAGGTACA	1173
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
DB	1174	GCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATGATGCTGAGGAAAGATC	1232
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
DB	1232	-----	1232
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
DB	1232	-----	1232
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
DB	1232	-----	1232
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
DB	1232	-----	1232
QY	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
DB	1232	-----	1232
QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
DB	1232	-----	1232
QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
DB	1232	-----	1232
QY	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
DB	1232	-----	1232
QY	501	GlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
DB	1233	-----CAAGTTGATGAAGTCAGAAAGG	1253
QY	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
DB	1254	CTGGTATATTTTGAAGGACCAAGACACTCCCTTTAGAGCATCACCCTGTACGTAGTCAGT	1313
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
DB	1314	TACGTAAATCTCGAGAGAGGTGACAAAGCTGACTGACCTGGCTACTCACATTCTTGCTGC	1373
QY	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
DB	1374	ATCAGTCAGCACTGTGACTTCTTATAAGTAAGTATAGTACCAAGAAATCCACTGT	1433
QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
DB	1434	GTGTCCCTTTTACAAGCTATCAAGTCTCAAGATGACCACTTGCATAAACAAGGAATTT	1493


```
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1494 TGGGCCACCAATTTTGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGCAAAATTTTC 1553
Qy 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1554 TCITTTGAAGTACTACTGGATTTATGTTATGATGGATGCTCTACAGCCCTCATGATCTA 1613
Qy 641 GlnProGlyLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL 660
Db 1614 CAGCCTGGAAGAAATATCTACTGTGCTTATATATGTTGCTCTCTCAGTGCACT 1673
Qy 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
Db 1674 TGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAG 1733
Qy 680 lYrYrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLeuPheGluG 700
Db 1734 GTTATGTGGTTGTAGTAGACACAGGGGATCTGTCCAGGGGCTTAAATTTGAAG 1793
Qy 700 lYAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnT 720
Db 1794 GCGCTTTAAATATAAATGGTCAATAGAAATTTGACGATCATGCTGGAAGGACTCCAAT 1853
Qy 720 YrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerT 740
Db 1854 ATCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATCCAGGCTGTGCT 1913
Qy 740 YrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 1914 ATGGAGGATACCTTCCCTCGATGGCATTAAATGACAGAGTCAAGATATCTTCAGGGTGTCTA 1973
Qy 760 leAlaGlyAlaProValThrLeuTyrPheThrAspThrGlyTyrThrGluArgTyrM 780
Db 1974 TTGCTGGGGCCCCGACGTCATCTGTGATCTTCTATGATACAGGATACACGGAAGCTTATA 2033
Qy 780 etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG 800
Db 2034 TGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCCATCGACGAG 2093
Qy 800 lYlYsPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheIleAspGluAsnV 820
Db 2094 AAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTATCATGGTTTCTCGATGAGAATG 2153
Qy 820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA 840
Db 2154 TCCATTTTGACATACAGTATATTTACTGAGTTTTTTTACTGAGGGCTGGAAGCCATATG 2213
Qy 840 spLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHist 860
Db 2214 ATTTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATT 2273
Qy 860 YrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuL 880
Db 2274 ATGAACATGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTTGCTGCTCTAA 2333
Qy 880 yValIle 882
Db 2334 AAGTGATA 2341
```

RESULT 9

```
US-09-976-674-4
; Sequence 4, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
```

```
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-4

Alignment Scores:
Pred. No.: 1-92e-315 Length: 2617
Score: 2870.00 Matches: 517
Percent Similarity: 77.5% Conservative: 134
Best Local Similarity: 61.5% Mismatches: 187
Query Match: 61.1% Indels: 2
DB: Gaps: 2
```

US-10-825-632-1 (1-982) x US-09-976-674-4 (1-2617)

```
Qy 35 PheTyrValGluArgTyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLys 54
Db 80 TTCAGGTGCAGAAACACCTCGTGGACGGGCTCCGGAGCATCATCCAGCGCGCCGCAAG 139
Qy 55 TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
Db 140 TACTCGGGCTTCATTGTCAACAAAGCGGCCCCACGACTTCCAGTTTGTGCAGAACCGGAT 199
Qy 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
Db 200 GAGTCTGGGGCCCCACTCCACCGCTCTACTACCTGGGAATGCCATATGGCAGCGGAGAG 259
Qy 95 AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu 114
Db 260 AACTCCCTCTCTACTCTGAGATTTCCAAAGAGGTCCGGAAGAGGCTCTGCTGCTCTG 319
Qy 115 SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
Db 320 TCCTGGAAGCAGATGCTGATCATTTTCCAGGCCACGCCCCACCATCTCTACTCTCGG 379
Qy 135 GluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp 154
Db 380 GAGAGAGAGCTGCTGAGGAGGCGGAACCCCTGGGGGTCTTCGGGCATCACCTCTCTAGAC 439
Qy 155 TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys 174
Db 440 TTCACAGCGAGAGTGGCTCTCTCTTCCAGGCCACGACCAACAGCCTCTTCCACTGTGCG 499
Qy 175 AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer 194
Db 500 GACGCGGCAAGAACCGCTTCATGGTGTCTCCCTATGAAACCGCTGGAAATCAAGACCCAG 559
Qy 195 CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe 214
Db 560 TGCTCAGGGCCCCGGATGAGACCCCAAAATCTGCCCTGCGACCCCTGCTCTCTCTCTTC 619
Qy 215 IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgArgLeuThr 234
Db 620 ATCAATAACAGCGACCTGTGGGTGGCCAAATCGAGACAGCGGAGGAGCGCGGCTGACC 679
Qy 235 TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr 254
Db 680 TTCCTGCCACCAAGGTTTATCCAATGTCTGGATGACCCCAAGTCTGCGGGTGTGGCCACC 739
Qy 255 PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGlu 274
Db 740 TTGCTCATACAGAGAGTTCGACCGCTTCACTGGGTACTGGTGTGCTGCCACACCTCC 799
Qy 275 ThrThrProSerGlyGly--LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer 293
Db 800 TGGGAAGTTTCAGAGGGCTTCAGAGCGCTGCGAATCTCTGTATGAGAAAGTGCATGAGTCC 859
Qy 294 GluValGluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPhe 313
```

Db 860 GAGGTGAGGTCAATCAGCTCCCTCTCTCGCGCTAGAGAAAGGACGCGACTCGTAT 919
Qy 314 ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMet 333
Db 920 CGGTACCCAGGACAGCAGCAAGATCCCAAGATTGCTTGAACCTGGCTGAGTCCAG 979
Qy 334 IleAspAlaGluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGlu 353
Db 980 ACTGACAGCCAGGCGAGATCGTCTCGACCCAGGAGAGGAGCTGGTGCACCCCTTCAGC 1039
Qy 354 IleLeuPheGluGlyValGluTyrIleAlaAsnAlaGlyTyrThrProGluGlyLysTyr 373
Db 1040 TCGCTGTTCGCAAGGTGGAGTACATCGCGGCGCGGTGACCCGGATGGCAATAC 1099
Qy 374 AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGluIleValLeuIleSerPro 393
Db 1100 GCCTGGCCATGTTCTGGACCGGCCAGAGTGGCTCCAGCTCGTCTCTCTCCCGCG 1159
Qy 394 GluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSerVal 413
Db 1160 GCCTGTTCATCCCGACACAGAGAAATGAGGACGCGCTAGCCTCTGCGAGAGCTGTC 1219
Qy 414 ProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIle 433
Db 1220 CCCAGGAATGTCAGCGCGTATGTGGTACGAGGAGGTCAACCAAGCTCTGGATCAATGTT 1279
Qy 434 HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
Db 1280 CATGACATCTTCTATCCCTTCCCGCAATCAGAGGGAGAGAGAGCTGCTGCTTCTCCGC 1339
Qy 453 AlaSerGluCysValThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGlu 472
Db 1340 GCCAATGAATGCAGACCGGCTTCTGCCAATTTGTACAAAGTCACCGCGCTTTTAAATCC 1399
Qy 473 SerLysTyrLysArgSerSerGlyLeuProAlaProSerAspPheLysCysProIle 492
Db 1400 CAGGCTACGATGAGTGAGTCCCTCAGCCCGCGGAGAGATGAATTTAAGTGCCCAT 1459
Qy 493 LysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn 512
Db 1460 AAGGAAGATTTGCTCTGACCGCGGTGATGGAGGTTTGGCGAGGACCGCTCCAAG 1519
Qy 513 IleGlnValAspGluValArgLeuValThrPheGluGlyThrLysAspSerProLeu 532
Db 1520 ATCTGGGTCAATGAGGAGACCAAGCTGTGTACTTCCAGGCGCACCAAGACACCGCGTG 1579
Qy 533 GluHisLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp 552
Db 1580 GAGCACCACTCTACGTGTGAGTATGAGCGCGCGGAGATCGTACGCTTCCACCG 1639
Qy 553 ArgGlyTyrSerHisSerCysIleSerGlnHisCysAspPhePheIleSerLysTyr 572
Db 1640 CCGCGCTTCTCCCATAGCTGTCTCATGAGCCAGAACTTCGACATGTTCTGCTACGCACTAC 1699
Qy 573 SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp 592
Db 1700 AGCAGCGTGAGCAGCGCGCTCGTGCACGCTCTACAAGCTGAGCGCGCGCCGACGAC 1759
Qy 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro 612
Db 1760 CCCTGCACAGCAGCGCGCTCTCTGGGTAGTAGTATGAGGAGGAGCGAGCTGCCCGCG 1819
Qy 613 AspTyrThrProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
Db 1820 GATTATGTTCTCTCAGAGATCTTCCATTTCCACAGCGCGCTCGGATGTCGGCTCTACGGC 1879
Qy 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysValTyrProThrValLeuPheIle 652
Db 1880 ATGATCTACAGGCCCGCGCTGTGAGCAGGAGAGACACCCACCGCTCTCTTTGTA 1939
Qy 653 TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArg 672

Db 1940 TATGAGGCGCCAGGTGAGCTGGTGAATAACTCTCTCAAAGGCATCAAGTACTTGGG 1999
Qy 673 LeuAsnThrLeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCys 692
Db 2000 CTCAACACATCGCTCCCTCGGCTACGCGTGGTGTGATTGACGCGAGGGCTCTGT 2059
Qy 693 HisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAsp 712
Db 2060 CAGCGAGGGCTTCGGTTCGAAGGGGCTCTGAATAACCAATGGCGAGTGGAGATCGAG 2119
Qy 713 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 732
Db 2120 GACCAGGTGAGGCGCTTCAGTTCTGTCGCGAGAGATATGGCTTTCATGACCTGAGCCGA 2179
Qy 733 ValGlyIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 752
Db 2180 GTTGCCATCCATGGCTGCTCTACGCGGGCTTCTCTCGCTCATGGGCTAATCCACAAG 2239
Qy 753 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 772
Db 2240 CCCAGGTGTTCAAGGTGGCCATCGCGGTGCGCGCTCACCGTCTGGATGGCTTACGAC 2299
Qy 773 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 792
Db 2300 ACAGGTACACTGACGCTACATGGAGCTCCCTGAGAACACACACGACGCTATGAGCG 2359
Qy 793 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 812
Db 2360 GTTTCGGTGGCGCTTCGACGTGGAGAGCTGCCCAATGAGCCCAACCGCTTGTCTTATCCTC 2419
Qy 813 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 832
Db 2420 CAGCGCTTCTTGAGAGAAACGTGCATTTTCCACAACTTCTCTCTCCCACTG 2479
Qy 833 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGluGluArgHisSerIleArg 852
Db 2480 ATCCAGGAGGAGAACTTACCAGTCCAGATCTACCCACGAGAGACACAGTATTGCG 2539
Qy 853 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 872
Db 2540 TGCCCCGAGTGGCGGAGCACTGAAGTCAGCTTGTGCACCTTCTACAGGATACTCTC 2599

RESULT 10
US-09-976-674-28
; Sequence 28, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 4219
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-976-674-28

Alignment Scores: 4,33e-315 Length: 4219
Pred. No.: 517
Score: 2870.00 Matches: 517
Percent Similarity: 77.5% Conservative: 134
Best Local Similarity: 61.5% Mismatches: 187
Query Match: 61.1% Indels: 2
DB: 3 Gaps: 2

QY	753	SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuThrIlePheTyrAsp	772	QY	115	SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg	134
DB	2596	CCCAGGTGTTCAAGGTGGCCATCGCGGGTGCCTGGGTCACCGTCTGGATCGCATCGAC	2655	DB	676	TCTTGGAGAGAGATGCTGGATCATTTCCAGGCCACGCCACCACCTGGGTCTACTCTCG	735
QY	773	ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrLeu	792	QY	135	GluGluLeuLeuAspGluArgLysArgIleGlyThrValGlyLysLeuAspTyrAsp	154
DB	2656	ACAGGGTACACTAGCGCTACATGGACGCTCCCTGGAGAACACACAGCAGCGGTATGAGCG	2715	DB	736	GAGGAGAGCTGCTGAGGAGCGGAACCGCTGGGGTCTTCGGCATCACCTCTCTACGAC	795
QY	793	GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu	812	QY	155	TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyLysLeuValLys	174
DB	2716	GTTTCCGTGGCCCTGCACGTGGAGAAAGCTGCCCAATGAGCCCAACCGCTTCTATCCTC	2775	DB	796	TTCACACGAGAGTGGCTCTTCTCTTCCAGGCCGACACAGAGCTCTTCACCTGCGCG	855
QY	813	HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu	832	QY	175	AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer	194
DB	2776	CACGGCTTCTGGACGAAACGTCACCTTTTCCACACAACTTCTCTGCTCTCCCACTG	2835	DB	856	GACGGCGCCCGATGGACCCCAAAATCTGCCCTGCGCAGCCCTGCTCTTCTCTCTC	915
QY	833	ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg	852	QY	195	CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe	214
DB	2836	ATCCGAGCAGGGAAACCTTACCAGCTCCAGATCTACCCCAACGAGAGACACAGATTGCG	2895	DB	916	TGCTCAGGGCCCGGATGGACCCCAAAATCTGCCCTGCGCAGCCCTGCTCTTCTCTCTC	975
QY	853	ValProGluSerGlyGluHisTyrGluLeuHisLeuHisTyrLeuGlnGluAsnLeu	872	QY	215	IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgLeuThr	234
DB	2896	TGCCCCGAGTCGGCGAGCACTATGAAGTCACGTTGCTGCACTTCTACAGGAATACCTC	2955	DB	976	ATCAATAACAGCAGCTGTGGTGGCCCAACATCGACACAGGCGAGAGCGGGCTGAGCC	1035
RESULT 11				QY	235	TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr	254
US-09-976-674-24				DB	1036	TTCTGCCACCAAGGTTTATCCAATGTCTGGATGACCCCAAGTCTCGGGTGTGGCCACC	1095
; Sequence 24, Application US/09976674				QY	255	PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGlu	274
; Patent No. 6844180				DB	1096	TTCTGTACATACAGGAAGATTCGACCGCTTCACTGGGTACTGTGGTGTGGCCACAGCTCC	1155
; GENERAL INFORMATION:				QY	275	ThrThrProSerGlyGly--LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer	293
; APPLICANT: Akinsanya, Karen				DB	1156	TGGGAAGGTTTACAGGGGCTCAAGACGCTCGGAATCTCTGTATGAGGAAGTCGATGAGTCC	1215
; APPLICANT: Riviere, Pierre				QY	294	GluValGluIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPhe	313
; APPLICANT: Junien, Jean-Louis				DB	1216	GAGGTGAGAGTCATTTCAGCTCCCTCTCTCGCTAGAGAAGGAAGACGAGCTCGTAT	1275
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV				QY	314	ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMet	333
; FILE REFERENCE: 706669				DB	1276	CGGTACCCGAGCAGGACGAGCAAGATCCCAAGATTCCTTGAATCGCTGAGTGTCCAG	1335
; CURRENT APPLICATION NUMBER: US/09/976,674				QY	334	IleAspAlaGluGlyArgIleAspValIleAspLysGluLeuIleGlnProPheGlu	353
; CURRENT FILING DATE: 2001-10-12				DB	1336	ACTGACAGCCAGGGCAAGATCGTCTCGACCCAGGAGGAAGGCTGTGTGACGCTTCAGC	1395
; PRIOR APPLICATION NUMBER: US 60/240,117				QY	354	IleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyr	373
; PRIOR FILING DATE: 2000-10-12				DB	1396	TGCTGTTCCTCCGAAAGTGGAGTACATCGCAGGCGCGGGTGGACCCCGGATGGCAATAC	1455
; NUMBER OF SEQ ID NOS: 61				QY	374	AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerPro	393
; SOFTWARE: PatentIn version 3.1				DB	1456	GCCTGGCCATGTTCTCTGACCGGCCCGCCAGCAGTGGCTCCAGCTCGCTCTCTCCCCCG	1515
; SEQ ID NO 24				QY	394	GluLeuPheIleProValGluAspAspValMetGluArgGlnArgLeuIleGluSerVal	413
; TYPE: DNA				DB	1516	GCCCTGTTTCATCCGAGCAGACAGAGATGAGGAGCAGCGGCTAGCCTCTGCCAGAGCTGC	1575
; ORGANISM: Homo sapiens				QY	414	ProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIle	433
US-09-976-674-24				DB	1576	CCGAGGAATGCCAGCGGTATGTGTGTACGAGGAGTCAACACGCTCTGGATCAATGTT	1635
Alignment Scores:				QY	434	HisAspIlePheHisValPheProGlnSerHis--GluGluGluIleGluPheIlePhe	452
Pred. No.:	4,48e-315	Length:	4302	DB	1636	CATGACATCTTATCTATCCCTTCCCAATCAGAGGAGGAGGACGAGCTCTGCTTCTCCGC	1695
Score:	2970.00	Matches:	517	QY	453	AlaSerGluCysIleThrGlyPheArgHisLeuTyrIleThrSerIleLeuLysGlu	472
Percent Similarity:	77.5%	Conservative:	134	DB	1696	GCCATGATGACAGACCGGCTCTGCAATTTGTACAAAGCAGCGCGTTTAAATATCC	1755
Best Local Similarity:	61.5%	Mismatches:	187	QY	473	SerLysTyrArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIle	492
Query Match:	61.1%	Indels:	2				
DB:	3	Gaps:	2				
US-10-825-632-1 (1-882) x US-09-976-674-24 (1-4302)							
QY	35	PheTyrValGluArgTyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLys	54				
DB	436	TTCAGGTGCAAGACACTGCTGGGACCGGCTCCGGAGCATCATCCAGCGCAGCCGACAG	495				
QY	55	TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp	74				
DB	496	TACTCGGGCTCATTTGTCAACAGCGCCCGCCACACTTCCAGTTTGTGCAAGACGAGAT	555				
QY	75	ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu	94				
DB	556	GAGTCCTGGGCGCCACTCCACCGCTCTACTACTTGGGAATGCCATATGCCAGCCGAGAG	615				
QY	95	AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu	114				
DB	616	AACTCCCTCTCTACTCTGAGATTCCCAAGAAGTCCCGAAGAGGCTCTGCTGCTCCTG	675				

```
Db 1756 CAGGGCTACGATTGGAGTGAGCCCTTCAGCCCGGGGAGAGATGAATTTAAGTGCCCAATT 1815
Qy LysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn 512
Db 1816 AAGGAAGAGATTGCTCTGACACACGGTGAATGGAGAGTTTGGCGAGGACGGCTCCAG 1875
Qy 513 IleGlnValAspGluValArgLeuValTyrPheGluGlyThrLysAspSerProLeu 532
Db 1876 ATCTGGGTCAATGAGGAGACCAAGCTGGTGTACTTCCAGGGCCACCAAGGACACGCGGTG 1935
Qy 533 GluHisIleLeuValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp 552
Db 1936 GAGCACCACCTCTACGTGGTGCAGCTATGAGCGCGCGCGGAGATCGTACGCTCACACG 1995
Qy 553 ArgGlyTyrSerHisSerCysValIleSerGlnHisCysAspPhePheIleSerLysTyr 572
Db 1996 CCCGGCTCTCCCATAGCTGCTCATGAGCCAGCACTTCGACATGTTGTCAGCCACTAC 2055
Qy 573 SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp 592
Db 2056 AGCAGCTGAGCAGCGCGCTGCTGTCAGCTCTCAAGCTGAGCGCGCGCGGACGAC 2115
Qy 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro 612
Db 2116 CCCCTGCACAGCAGCCCGCTTCTGGCTAGCATGATGAGGCGACGAGCTGCCCGCCG 2175
Qy 613 AspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
Db 2176 GATTATGTTCTCCAGAGATCTTCCATTTCCACACCGCTCGGATGTGCGGCTCTACGGC 2235
Qy 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIle 652
Db 2236 ATGATCTACAGCCCGCCAGCGCTGCGACGAGGAGAGACACCCACCGCTCTCTTGTGA 2295
Qy 653 TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArg 672
Db 2296 TATGGAGGCGCCCGAGGTGCAGCTGGTGAATACTCTTCAAAGGCATCAAGTACTTGGCG 2355
Qy 673 LeuAsnThrLeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCys 692
Db 2356 CTCAACACACTGGCGCTCCCTGGGCTACGCGTGGTGTGATTGACGCGGAGGGGCTCTGT 2415
Qy 693 HisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAsp 712
Db 2416 CAGCAGGGCTTCGGTTCGAAAGGGCCCTGAAAAACCAATGGGCCAGGTGGAGATCGAG 2475
Qy 713 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 732
Db 2476 GACCAGGTGGAGGGCTGCAGTTGCTGGCGGAGAGATGATGGCTTCATCGACCTGAGCCGA 2535
Qy 733 ValGlyIleHisGlyTyrSerTyrGlyTyrLysSerLeuMetAlaLeuMetGlnArg 752
Db 2536 GTTGCCATTCATGCTGCTGCTACGGGGGCTTCTCTCGCTCATGCGGGCTAATCCACCAAG 2595
Qy 753 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 772
Db 2596 CCCAGGTTCAAGAGTGGCCATCGCGGGTGCCCGGCTCACCGTCTGGATGGCTTACGAC 2655
Qy 773 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 792
Db 2656 ACAGGTTACATGAGCGCTACATGAGCGTCCCTGAGNACACACCGACGCGCTATGAGCG 2715
Qy 793 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 812
Db 2716 GGTTCGGTGGCCCTGCAGTGGAGAGCTGCCCAATGAGCCCAACCGCTTGTGTATCTCTC 2775
Qy 813 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 832
Db 2776 CACGGCTTCCTGGAGCGAAACGTCACTTTTTCACACAAACTTCTCGTCTCCCACTG 2835
Qy 833 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 852
```

```
Db 2836 ATCCGAGCAGGAAACCTTACCAGCTCCAGATCTACCCCAACGAGACACAGTATTTCG 2895
Qy 853 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 872
Db 2896 TGCCCCGAGTCCGGGAGCACTATGAAGTCAAGTTCGTCACCTTTCTACAGGAATACCTC 2955

RESULT 12
US-09-976-674-36
; Sequence 36, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Oi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 4180
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-36

Alignment Scores:
Pred. No.: 1,82e-309 Length: 4180
Score: 2820.50 Matches: 510
Percent Similarity: 76.4% Conservative: 132
Best Local Similarity: 60.7% Mismatches: 183
Query Match: 60.0% Indels: 15
DB: 3 Gaps: 3

US-10-825-632-1 (1-882) x US-09-976-674-36 (1-4180)
Qy 35 PheTyrValGluArgTyrSerTyrSerGlnLeuLysLysLeuLeuAlaAspThrArgLys 54
Db 436 TTCAGGTGCAGAGCACTCTGGGACGGGTCCGGAGCATCATCCACGGCAGCCGCAAG 495
Qy 55 TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
Db 496 TACTCGGGCTCATTTGTCACAAAGCGCGCCACGACTTCCAGTTTGTGCAAGACGAT 555
Qy 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
Db 556 GAGTCTGGGGCCCACTCCCGCTCTACTACCTGGGAATGCCATATGCGACGCCGAGAG 615
Qy 95 AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu 114
Db 616 AACTCCCTCTCTACTCTGAGATCCCAAGAGGTCCGGAAGAGGCTCTGCTGCTCTG 675
Qy 115 SerTyrLysProLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
Db 676 TCTTGAAGCAGATCTGGATCATTTCCAGCCACGCGCCACCATGGGTCTACTCTCGG 735
Qy 135 GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp 154
Db 736 GAGGAGGAGCTCTGAGGAGCGGAAACGCTGGGGGTCTTTCGGCATCACCTCTCTACGAC 795
Qy 155 TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys 174
Db 796 TTCACAGCAGAGAGTGGGCTCTTCTCTTCCAGGCCAGCAACACCTCTTCCACTGCCG 855
Qy 175 AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer 194
Db 856 GACGGGGCAAGAACGGCTTTCATGTGTCCCTATGAACCGCTGGAAATCAACAGACCCAG 915
Qy 195 CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe 214
```

Db 916 TGCTCAGGCGCCGATGACCCCAAAATCTGCCCTGCCAGCTCTTCTCTCTC 975
Qy 215 IleHisSerAenAspIleThrPheSerAenIleValThrArgGluGluArgLeuThr 234
Db 976 ATCAATAACAGCAGCTGTGGTGGTCAACATCGAGACAGCGAGGAGCGCGCTGACC 1035
Qy 235 TyrValHisAenGluLeuAlaAenMetGluGluAenAlaArgSerAlaGluValAlaThr 254
Db 1036 TTCTGCCACCAAGGTTTATCCAAATGCTCTGGATGACCCCAAGTCTGGGGGTGTGCCACC 1095
Qy 255 PheValLeuGluGluPheAspArgTyrSerGlyTyrTyrPheValPheAlaGlu 274
Db 1096 TTCTGTCATACAGAGAGTTCCAGCCGCTTCACTGGGTACTGGTGGTCCCCACAGCTCC 1155
Qy 275 ThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAenAspGluSer 293
Db 1156 TGGGAAGGTTCCAGAGGCTCAAGACGCTGCAGATCTCTGTATGAGGAAGTGCATGATGCC 1215
Qy 294 GluValGluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPhe 313
Db 1216 GAGGTGAGGTCAATTACGTCCTCTCTGCGCTAGAGAAAGGAGACCGGACTCGTAT 1275
Qy 314 ArgTyrProLysThrGlyThrAlaAenProLysValThrPheLysMetSerGluIleMet 333
Db 1276 CGGTACCCAGGACAGCAGCAAGAAATCCCAAGATTGCTTGAAACTCGCTGAGTCCOAG 1335
Qy 334 IleAspAlaGluGluArgIleIleAspValIleAspLysGluLeuIleGluProPheGlu 353
Db 1336 ACTGACAGCCAGGCGAAGATCGTCTCGACCCAGGAGAGGAGCTGGTGCAGCCCTTCAGC 1395
Qy 354 IleLeuPheGluGluValGluTyrIleAlaArgAlaGlyTyrThrProGluGluLysTyr 373
Db 1396 TCGCTGTTCCGAGGTGGAGTACATCGCCAGGCGCGGTGGACCCGGATGGCAATAC 1455
Qy 374 AlaThrPheIleLeuAspArgSerGlnThrArgLeuGluIleValLeuIleSerPro 393
Db 1456 GCTTGGGCGCATGTCTTGACCGCCGCCAGCGTCCAGCTCGTCTCTCTCCCGCG 1515
Qy 394 GluLeuPheIleProValGluAspValMetGluArgGluArgLeuIleGluSerVal 413
Db 1516 GCCCTGTTCATCCGACACAGAGATGAGGACGCGGTAGCCTCGCCAGGCTGTC 1575
Qy 414 ProAspSerValThrProLeuIleIleTyrGluGluThrAspIleThrIleAsnIle 433
Db 1576 CCCAGGAATGCCAGCGGTATGTGGTGTACGAGGAGGTCAACACGCTCTGGATCAATGTT 1635
Qy 434 HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
Db 1636 CATGACATCTTCTATCCCTTCCCCCAATCAGAGGGAGAGGAGCGTCTGCTTCTCCGC 1695
Qy 453 AlaSerGluCysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGlu 472
Db 1696 GCCAATGAATCAAGACCGGCTTCTGCCATTTGTCAAAAGTCAACCGCGCTTTTAAATCC 1755
Qy 473 SerLysTyrLysArgSerSerGlyLeuProAlaProSerAspPheLysCysProIle 492
Db 1756 CAGGCTACGATCGAGTGAGCCCTTCAGCCCGGGGAAGATGAAATTTAAGTGCCCATTT 1815
Qy 493 LysGluGluIleAlaIleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAen 512
Db 1816 AAGGAAGATGCTCTGACCCGCGGTGATGGAGGTTTGGCGAGGACCGCTCC--- 1872
Qy 513 IleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeu 532
Db 1873 -----AAGGGCACCAAGGACAGCGCGCTG 1896
Qy 533 GluHisHisLeuTyrValValSerTyrValAsnProGlyValThrArgLeuThrAsp 552
Db 1897 GAGCACCACTCTACGTGGTCACTATGAGCGCGCGCGGAGATCGTACGCTTCACACG 1956
Qy 553 ArgGlyTyrSerHisSerCysCysIleSerGlnHisCysAspPhePheLysTyr 572
Db 1957 CCCGGCTTCTCCCATAGTGTCTCATGAGCCAGAACTTCGATGTTCTGTCGACCACTAC 2016

Qy 573 SerAenGlnLysAenProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp 592
Db 2017 AGCAGCTGAGCAGCGCCGCTCGTGCAGCTTACAGCTGAGCGCGCCGACGACGAC 2076
Qy 593 ProThrCysLysThrLysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuPro 612
Db 2077 CCCCTGCACAGCAGCGCCGCTTCTGGGTAGCATGTGAGGCGACGAGCTGCCCGCG 2136
Qy 613 AspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
Db 2137 GATTATGTTCTCCAGAGATCTTCCATTTCCACACGCGCTCGGATGTGCGGCTCTACGGC 2196
Qy 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIle 652
Db 2197 ATGATCTACAAGCCCAAGCTTTCAGCCAGGAGGAAGACCCCAACCGCTCTCTTTGTA 2256
Qy 653 TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArg 672
Db 2257 TATGGAGCGCCCAAGGTGCAGCTGGTGAATACTCTTCAAAGGCATCAAGTACTTGGCG 2316
Qy 673 LeuAenThrLeuAlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCys 692
Db 2317 CTCACACACTGGCTCTCTGGGTAGCGCGCTGTGTGATTGACGCGAGGGCTCTCTGT 2376
Qy 693 HisArgGlyLysLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAsp 712
Db 2377 CAGCAGGCGCTTCGTTTCGAAGGGGCGCTGAAACCAATGGCCGAGGTGGAGATCGAG 2436
Qy 713 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 732
Db 2437 GACCAGGTGGAGGCGCTGCAGTTCGTGGCGGAGAAAGTATGGCTTCATCGACCTGAGCCGA 2496
Qy 733 ValGlyIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 752
Db 2497 GTTGCCATCTCAAGTGGCTCTCTACGGGGGCTCTCTCGCTCATGGGGCTATCCCAAG 2556
Qy 753 SerAspIlePheArgValAlaIleAlaGlyValAlaProValThrLeuThrIlePheTyrAsp 772
Db 2557 CCCAGAGTGTTCAAAGTGGCCATCGCGGGTGCCTCGCTCAACCGCTCTGGATGGCGTACGAC 2616
Qy 773 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 792
Db 2617 ACAGGGTACACTGAGCGCTACATGAGCGTCCCTGAGAAACAACGACGCGCTATGAGGCG 2676
Qy 793 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeu 812
Db 2677 GGTTCGTGGCGCTGCAGCTGAGAGCTGCCCAATGAGCCCAACCGCTTGTCTTATCTTC 2736
Qy 813 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 832
Db 2737 CACGGCTTCTGGAGGAACGTGCATTTTCCACCAAACTTCTCTCGTCTCCCACTG 2796
Qy 833 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 852
Db 2797 ATCCGAGCAGGGAAACCTTACAGCTCCAGATCTACCCCAACGAGAGACAGATATTCG 2856
Qy 853 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeu 872
Db 2857 TGCCCGAGTGGCGGAGCACTATGAAGTCACTGCTGTGCACTTTTACAGGAATACCTC 2916

RESULT 13
US-09-976-674-34
; Sequence 34, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674

Db 1456 GCCTGGCCCATGTTCTCGACCGCCCGCCAGAGTGGCTCCAGCTCGCTCTCTCCCTCCCG 1515
Qy 394 GluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSerVal 413
Db 1516 GCCTGTTTCATCCCGAGCAGAGAAATGAGGAGCAGCGCTAGCCTCTGCGAGAGTGTG 1575
Qy 414 ProAspSerValThrProLeuIleLeuThrGluGluThrThrAspIleThrPheLeu 433
Db 1576 CCACGAAATCTCCAGCGCTATGTTGTTGACGAGGAGGTACCAACGCTCGATCAATGTT 1635
Qy 434 HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
Db 1636 CATGACATCTTCTATCCTTCCCTCCCAATCAGAGGAGAGGAGCTCTGCTTCTCCCG 1695
Qy 453 AlaSerGluCysAlaThrGlyPheArgHisLeuThrLysIleThrSerIleLeuLysGlu 472
Db 1696 GCCAATGAATGCAAGACCGGCTCTGCCATTGTTGACAAAGTCACCGCGGTTTTAAATCC 1755
Qy 473 SerLysThrLysArgSerSerGlyLeuProAlaProSerAspPheLysCysProIle 492
Db 1756 CAGGCTACGATTGGAGTGGAGCTTTCAGCCCGCGGGAAGATGAATTAAGTGGCCCAT 1815
Qy 493 LysGluGluIleAlaIleThrSerGlyGluThrGluValLeuGlyArgHisGlySerAsn 512
Db 1816 AAGGAGAGATTCCTGACCAAGCGGTGAATGGGAGGTTTGGCGGAGCAGCGCTCCAAG 1875
Qy 513 IleGlnValAspGluValArgArgLeuValThrPheGluGlyThrLysAspSerProLeu 532
Db 1876 ATCTGGGTCAATGAGGAGCAAGCTGTTGTTACTTCCAGGCGCCAGGACACCGCGCTG 1935
Qy 533 GluHisLeuThrValValSerThrValAsnProGlyGluValThrArgLeuThrAsp 552
Db 1936 GAGCACCACTCTAGCTGGTGGTACGTATGAGCGCGCGCGAGATCGTACGCGCTCACCACG 1995
Qy 553 ArgGlyThrSerHisSerCysLysIleSerGlnHisCysAspPheIleSerLysThr 572
Db 1996 CCCGGCTTCTCCATAGCTCTCTCCATGAGCCAGAACTTCGACATGTTTCGTGAGCCATC 2055
Qy 573 SerAsnGlnLysAsnProHisCysValSerLeuThrLysLeuSerSerProGluAsp 592
Db 2056 AGCAGCTGAGCAGCGCGCTCGTGGTGCAGCTCTACAGCTGAGCGGCGCCGAGCAGCAG 2115
Qy 593 ProThrCysLysThrLysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuPro 612
Db 2116 CCCCTGCACAGCAGCGCGCTCTCTGGCTAGCATGATGAGGAGCAGCGCTGCGCCCG 2175
Qy 613 AspThrThrProGluIlePheSerPheGluSerThrThrGlyPheThrLeuThrGly 632
Db 2176 GATTATGTTCTCCAGAGATCTTCCATTTCACACCGCGCTCGGATGTGCGGCTCTACGCG 2235
Qy 633 MetLeuThrLysProHisAspLeuGlnProGlyLysThrProThrValLeuPheIle 652
Db 2236 ATGATCTACAGCCCGCTTTCAGCGAGGAGGAGGAGCAGCCCGCTCTCTTTGTA 2295
Qy 653 TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysThrPheArg 672
Db 2296 TATGAGGCGCCCGAGGTGCGCTGTTGTAATACCTCTTCAAGGATCACTACTTGGG 2355
Qy 673 LeuAsnThrLeuAlaSerLeuGlyThrValValValValIleAspAsnArgGlySerCys 692
Db 2356 CTCAACACACTGGGCTCTCCCTGGGCTACGCGGTGTTGTTGATTGAGCGGAGGCGCTCTGT 2415
Qy 693 HisArgGlyLeuLysPheGluGlyAlaPheLysThrLysMetGlyGlnIleGluLeuAsp 712
Db 2416 CAGGAGGCGCTTCGTTTCGAGGGGCGCTTGAATAACCAATGGGCGAGGTGGAGATCGAG 2475
Qy 713 AspGlnValGluGlyLeuGlnThrLeuAlaSerArgThrAspPheIleAspLeuAspArg 732
Db 2476 GACCAAGTGGAGGCGCTTCGAGTTGCGGCGGAGAGTATGGCTTCATGACCTGAGCGGA 2535
Qy 733 ValGlyIleHisGlyThrSerThrGlyThrLeuSerLeuMetAlaLeuMetGlnArg 752
Db 2536 GTTGCCATCATGGCTCTCTACGGGGGCTTCTCTCTGCTCATGGGCTAATCCACAAG 2595

Qy 753 SerAspIlePhe-ArgValAlaIleAlaGlyAlaProValThrLeuThrPheThrAs 772
Db 2596 CCCACAGTGTTCAGG----- 2611
Qy 772 pThrGlyThrThrGluArgThrMetGlyHisProAspGlnAsnGluGlnGlyThrLe 792
Db 2611 ----- 2611
Qy 792 uGlySerValAlaMetGlnAlaGlyPheProSerGluProAsnArgLeuLeuLe 812
Db 2612 -----CCCAACCGCTTGTCTTATCCT 2631
Qy 812 uHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLe 832
Db 2632 CCACGGCTTCTCGAGCGAAGAGCTGCACCTTTTCCACACAACTTCTCGTCTCCCACT 2691
Qy 832 uValArgAlaGlyLysProThrAspLeuGlnIleThrProGlnGluArgHisSerIleAr 852
Db 2692 GATCCGAGCAGGAGAAACCTTACAGCTCCAGATCTACCCCAACGAGACACAGTATTCG 2751
Qy 852 gValProGluSerGlyGluHisThrGluLeuHisLeuLeuHisThrLeuGlnGluAsnLe 872
Db 2752 CTGCCCCGAGTCCGGCGAGCAGCTATGAGTCACTGCTGCTTCTTACAGGAATACCT 2811
Qy 872 u 872
Db 2812 C 2812

RESULT 15

US-09-976-674-30
; Sequence 30, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-30

Alignment Scores:
Pred. No.: 5,91e-290 Length: 4159
Score: 2649.00 Matches: 486
Percent Similarity: 72.9% Conservative: 127
Best Local Similarity: 57.8% Mismatches: 177
Query Match: 56.4% Indels: 51
DB: 3 Gaps: 3

US-10-825-632-1 (1-882) x US-09-976-674-30 (1-4159)

Qy 35 PheThrValGluArgThrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLys 54
Db 436 TTCAGGTGAGAGAGCACTCGTGGGACGGCTCCGGAGCATCATCCACGGCAGCGCAG 495
Qy 55 ThrHisGlyThrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
Db 496 TACTCGGCGCTCATTTGCAACAGCGCCCGCCAGCTTCCAGTTTGTGTCAGAACGCGAT 555
Qy 75 ProAspGlyProHisSerAspArgIleThrThrLeuAlaMetSerGlyGluAsnArgGlu 94
Db 556 GAGTCTGGGCGCCCACTCCCAACCGCTTACTACTCTGGGAATGCCATATGGCAGCGCAGAG 615

QY	95	AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu	114	QY	453	AlaSerGluCysIysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGlu	472
Db	616	AACTCCCTCTCTACTCTAGATTCCCAAGAGGTCCGGAAGAGGCTCTGCTGCTCTG	675	Db	1696	GCCATGATGCAAGACCGCTTCGCCATTTGTACAAAGTACACCGCGTTTAAATCC	1755
QY	115	SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg	134	QY	473	SerLysTyrLysArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIle	492
Db	676	TCCTGGAGACAGATGCTGGATCTTCAGGCGCACGCCCCACCATGGGGTCTACTCTCGG	735	Db	1756	CAGGGCTACGATGGAGTGGAGCCCTTCAGCCCGGGAAGATGAATTAAGTCCCAT	1815
QY	135	GluGluGluLeuLeuArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp	154	QY	493	LysGluGluLeuAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn	512
Db	736	GAGGAGAGCTGCTAGGGAGCGGAACGCTGGGGTCTTCGGCATCACCTCTACGAC	795	Db	1816	AAGAAAGAGATGTCTGACACCGGTGAATGGAGGTTTTGGCAGCACCGCTCCAAG	1875
QY	155	TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys	174	QY	513	IleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeu	532
Db	796	TTCCACAGCAGAGTGGCTCTCTCTTCAGGCCAGCAACAGCTCTTCACCTGCCGCG	855	Db	1876	ATCTGGGTCAATGAGGAGACCAAGCTGGTGATCTTCAGGGCACCAGACACGCGCTG	1935
QY	175	AspGlyGlyProGlnGlyPheThrGlnProLeuArgProAsnLeuValGluThrSer	194	QY	533	GluHisHisLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp	552
Db	856	GACGGCGCAAGACCGCTTCATGGTGTCCCTATGAACCTCGAATCAGACCCAG	915	Db	1936	GAGCACCACTCTACGTGGTCAGCTATGAGCGCGCGCGGAGATCGTACGCTCACACG	1995
QY	195	CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe	214	QY	553	ArgGlyTyrSerHisSerCysValIleSerGlnHisCysAspPhePheIleSerLysTyr	572
Db	916	TGCTCAGGCGCCCGGATGGACCCCAAAATCTGCCCTGCGACCCCTGCTTCTCTTC	975	Db	1996	CCGGCTTCTCCCATAGCTGCTCCATGAGCCAGAACTTCGACATGTTCTGACCCACTAC	2055
QY	215	IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgLeuThr	234	QY	573	SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp	592
Db	976	ATCAATAACAGCGACTGTGGTGGCCCAACATCGACACAGCGGAGGCGGCTGACC	1035	Db	2056	AGCAGCGTGACACGCCGCCCTGCTGTCAGCTCTCAAGCTGAGCGGCCCCGACGAC	2115
QY	235	TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr	254	QY	593	ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro	612
Db	1036	TTCTGCCACCAAGGTTTATCCAATGCTCTGGATGATGCCCAAGTCTCGGGTGTGGCCACC	1095	Db	2116	CCCTGTGCACAGACCCCGCTTCTGGCTTAGCATGATGAGGAGCAGCAGCTGCCCGC	2175
QY	255	PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpCysProLysAlaGlu	274	QY	613	AspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly	632
Db	1096	TTCTCATACAGAGAGTTTCGACCGCTTCACTGGGTACTGGTGTGCCCCACAGCCTCC	1155	Db	2176	GATTATGTCTCCAGAGATCTTCATTTCCACGCGCTCGGATGCGGCTCTACGCG	2235
QY	275	ThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer	293	QY	633	MetLeuTyrLysProHisAspLeuGlnProGlyLysIysTyrProThrValLeuPheIle	652
Db	1156	TGGGAAGTTTCAGGGGCTCAAGCGCTCGCAATCTCTGATGAGGAAGTCGATGATCC	1215	Db	2236	ATGATCTAAGCCGCCACCGCTTCGACCGAGGAGAGACCCCACTCTCTTTGTA	2295
QY	294	GluValGluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAspPhe	313	QY	653	TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArg	672
Db	1216	GAGTGGAGGTCAITTCAGTCCCTCTCTCGCTAGAGAAAGGAGCGGACTCGTAT	1275	Db	2296	TATGAGGCCCCAGTGCAGCTGGTGAATACTCTCAAGGCATCAAGTACTTGC	2355
QY	314	ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMet	333	QY	673	LeuAsnThrLeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCys	692
Db	1276	CGGTACCCAGACAGGACGAGCAAGATCCCAAGATCCCTTGAATGCTGATGCTCCAG	1335	Db	2356	CTCAACACACTGGCTCCCTGGCTACCGCTGGTGTGTGATTACGCGAGGGCTCTGT	2415
QY	334	IleAspAlaGluGlyArgIleLeuAspValIleAspLysGluIleGlnProPheGlu	353	QY	693	HisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAsp	712
Db	1336	ACTGACAGCCAGGCGAAGATCGTCTCGACCCAGGAGAGGAGCTGGTGCGCCCTTCAGC	1395	Db	2416	CAGCGAGGGCTTCGGTTCCGAAGGGGCCCTGAAAAACCAATGGGCCAGGTGGAGATCGAG	2475
QY	354	IleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyr	373	QY	713	AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg	732
Db	1396	TCGCTGTTCGGAAGTGGAGTACATCGCCAGGGCCGGGTGGACCCGGGATGGCAATAC	1455	Db	2476	GACCAAGTGGAGGCGCTCGAGTTCGTGGCCGAGAAGTATGCTTCATCGACCTGAGCCGA	2535
QY	374	AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerPro	393	QY	733	ValGlyIleHisGlyTrpSerTyrGlyGlyTyrSerLeuSerLeuMetAlaLeuMetGlnArg	752
Db	1456	GCCTGGCCATGTCTCGACCGGCCCCAGCAGTGGCTCCAGCTCGTCTCTCCCTCC	1515	Db	2536	GTTGCCATTCATGGCTGGTCTCTACGGGGCTCTCTCTCGCTCATGGGGCTTAATCCACAAG	2595
QY	394	GluLeuPheIleProValGluAspValMetGlnArgGlnArgLeuIleGluSerVal	413	QY	753	SerAspIlePhe---ArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAs	772
Db	1516	GCCCTGTTCATCCCGACAGAGAATGAGGACGAGCGGCTAGCCTCTGCGCAGAGCTGTC	1575	Db	2596	CCCCAGGTGTTCAGS-----	2611
QY	414	ProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIle	433	QY	772	pThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLe	792
Db	1576	CCNAGGATGTCAGCGGTATGTGGTGTACGAGGAGGTCCACCAAGCTCGATCAATGTT	1635	Db	2611	-----	2611
QY	434	HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe	452	QY	792	uGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu	812
Db	1636	CATGACATCTCTCTCCCTTCCCCCAATCAGAGGAGGAGGAGCAGCTCTGCTTCTCCG	1695	Db	2612	-----CCCAACCGCTGCTTATCCT	2631
Db				QY	812	uHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLe	832

Db	2632	CCACGGCTTCCTGGACGAAACGTGCACCTTTTCCACACAACTTCTCGTCTCCAACT	2691
QY	832	uValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleAr	852
Db	2692	GATCCGACGAGGAAACCTTACACAGCTCCAGATCTACCCCAACGAGACACAGTATTCG	2751
QY	852	gValProGluSerGlyGluHisTyrGluHisLeuLeuHisTyrLeuGlnGluAsnLe	872
Db	2752	CTGCCCCGAGTCGGGCGAGCAGTATGAAGTCAGTTGCTGCACCTTTCTACAGGAATACCT	2811
QY	872	u 872	
Db	2812	C 2812	

Search completed: April 15, 2006, 01:02:11
Job time : 495 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 01:02:29 ; Search time 1458 Seconds
(without alignments)
5002.463 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMETEQLGVSEIFETADC.....HLLHYLQENLGSRIALAKVI 882

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL=frame-p2n.model -DEV=xlp
-Q-/abs/ABSSWEB spool/US10825632/runat 14042006 092611 4971/app query.fasta_1
-DB=Published Applications NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05p
-USER=US10825632 @CGN 1 1 1549 @runat 14042006 092611 4971 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA_Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	100.0	2649	6	US-10-054-776-1
2	4700	100.0	2649	6	US-10-170-789-39
3	4700	100.0	2671	9	US-09-976-674-2
4	4700	100.0	2671	9	US-10-982-512-2
5	4700	100.0	3106	7	US-10-311-035-30
6	4700	100.0	3120	8	US-10-415-122-5
7	4700	100.0	3120	8	US-10-825-632-2

8	4700	100.0	3143	6	US-10-170-789-37	Sequence 37, Appl
9	4680	99.6	4829	3	US-09-976-674-12	Sequence 12, Appl
10	4680	99.6	4829	9	US-10-982-512-12	Sequence 12, Appl
11	4385.5	93.3	4685	3	US-09-976-674-22	Sequence 22, Appl
12	4385.5	93.3	4685	9	US-10-982-512-22	Sequence 22, Appl
13	4385	93.3	4676	3	US-09-976-674-20	Sequence 20, Appl
14	4385	93.3	4676	9	US-10-982-512-20	Sequence 20, Appl
15	4118	87.6	2830	3	US-10-956-157-2177	Sequence 2177, Ap
16	4092.5	87.1	4523	3	US-09-976-674-8	Sequence 8, Appli
17	4092.5	87.1	4523	9	US-10-982-512-8	Sequence 8, Appli
18	3970.5	84.5	2510	7	US-10-275-505-16	Sequence 16, Appl
19	3970.5	84.5	2510	10	US-11-140-224-16	Sequence 16, Appl
20	3661.5	77.9	4309	3	US-09-976-674-14	Sequence 14, Appl
21	3661.5	77.9	4309	9	US-10-982-512-14	Sequence 14, Appl
22	2870	61.1	2617	3	US-09-976-674-4	Sequence 4, Appli
23	2870	61.1	2617	9	US-10-982-512-4	Sequence 4, Appli
24	2870	61.1	3716	9	US-10-433-757-30	Sequence 30, Appl
25	2870	61.1	4219	3	US-09-976-674-28	Sequence 28, Appl
26	2870	61.1	4219	9	US-10-982-512-28	Sequence 28, Appl
27	2870	61.1	4302	3	US-09-976-674-24	Sequence 24, Appl
28	2870	61.1	4302	9	US-10-982-512-24	Sequence 24, Appl
29	2863	60.9	3000	7	US-10-415-122-1	Sequence 1, Appli
30	2862	60.9	2660	7	US-10-072-012-225	Sequence 225, App
31	2862	60.9	2660	7	US-10-072-012-225	Sequence 225, App
32	2835	60.3	2495	7	US-10-415-122-8	Sequence 8, Appli
33	2833	60.3	3287	7	US-10-415-122-3	Sequence 3, Appli
34	2820.5	60.0	4180	3	US-09-976-674-36	Sequence 36, Appl
35	2820.5	60.0	4180	9	US-10-982-512-36	Sequence 36, Appl
36	2820.5	60.0	4263	3	US-09-976-674-34	Sequence 34, Appl
37	2820.5	60.0	4263	9	US-10-982-512-34	Sequence 34, Appl
38	2649	56.4	4076	3	US-09-976-674-32	Sequence 32, Appl
39	2649	56.4	4076	9	US-10-982-512-32	Sequence 32, Appl
40	2649	56.4	4159	3	US-09-976-674-30	Sequence 30, Appl
41	2649	56.4	4159	9	US-10-982-512-30	Sequence 30, Appl
42	2638	55.3	2801	5	US-10-098-841-100	Sequence 100, App
43	2599.5	55.3	4037	3	US-09-976-674-40	Sequence 40, Appl
44	2599.5	55.3	4037	9	US-10-982-512-40	Sequence 40, Appl
45	2599.5	55.3	4120	3	US-09-976-674-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-10-054-776-1
; Sequence 1, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2649)
US-10-054-776-1

Alignment Scores:			
Pred. No.:	0	4700.00	Length: 2649
Score:			Matches: 882
Percent Similarity:	100.0%		Conservative: 0
Best Local Similarity:	100.0%		Mismatches: 0
Query Match:	6		Indels: 0
DB:			Gaps: 0

US-10-825-632-1 (1-882) x US-10-054-776-1 (1-2649)

QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 1 ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGGTGGATATTTGAAACATGGGACTGT 60
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 61 GAGAGAAATATTGAATCACAGATCGGCCCTAAATTTGAGGCCCTTTTATATTTGAGCGGTAT 120
QY 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 121 TCCTGGAGTCAGCTTAAAGAGCTGCTGGCGATACCAGAAATATATCGGCTACATGATG 180
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 181 GCTAAGGCACCATGATATTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA 240
QY 81 AspArgIleTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 241 GACAGAATCTATTACCTTGGCATGCTCGGTGAGAACAGAGAAATATACACTGTTTATTCT 300
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 301 GAAATTCCTCAAAATATCAATAGACGACAGCTCTTAATGCTCTCTTGGAAAGCCTCTTTG 360
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 361 GATCTTTTTCAGGCAACACTGGACTATGGAAATGTAATCTCGAGAAAGAAACTATTAAAGA 420
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 421 GAAGAAAAACCATTTGGNACNGTCGGNAATTTGCTTTCAGATATATACCAAGGAAGTGA 480
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 481 ACATTTCTGTTTCAAGCCGGTAGTGGNAATTTACGTAAAGATGGAGGCCACAGGA 540
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 541 TTTACGCAACAACTTTTAAGCCCAATCTAGTGGAACTAGTGTGCCCAACATACGGATG 600
QY 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
DB 601 GATCCAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 660
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
DB 661 TGGATATCTAAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACATGAGCTA 720
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 721 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAA 780
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 781 TTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGATGGTGT 840
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
DB 841 AAAATCTTACAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTTCATGTT 900
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 901 ACATCCCTATGTTTGGAAACAGAGGGCAGATTCTTCGTTATCTCTTAAACAGGTACA 960
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlyArgIle 340
DB 961 GCAAATCTTAAAGTCATTTTAAAGATGTCAGAAATATGATGTGATGCTGAAGGAAGATC 1020
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlyValGlu 360
DB 1021 ATAGATGTCATAGATAAGGAACATTAATCAACCTTTTGGATTTCTATTGGAAGGATTTGA 1080

QY 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1081 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCTACTAGAT 1140
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1141 CGCTCCAGACTCGCTTCGAGATAGTGTGATCTCACTGAATATTATTATCCAGTAGAA 1200
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
DB 1201 GATGATGTTATGGAAAGCGAGACTCATTTGATCGTCACTGCTGCTGAGCCACTA 1260
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1261 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTT 1320
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1321 CCCCAAGTCCAGAAAGAGAAATTTGAGTTTATTTTTTGCCTCTGAATGCAAAACAGGTTTC 1380
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGly 480
DB 1381 CGTCATTTTATACAAATTTACATCTATTATTTAAAGAAAGCAAAATATATAACGATCCAGTGT 1440
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
DB 1441 GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAAAGAGAGATAGCAATTTACAGT 1500
QY 501 GlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
DB 1501 GGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAG 1560
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
DB 1561 CTGGTATATTTTGAAGCCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCACT 1620
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisCysCys 560
DB 1621 TACGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGC 1680
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1681 ATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAAGATCCCACTGT 1740
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1741 GGTCCCTTTTCAAGCTATCAAGTCTTGAAGATGACCAACTTGCACAAACAAAGGAATTT 1800
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
DB 1801 TGGGCCACCATTTTGGATTTACAGCAGGTCTCTTCTCTGACTATACTCTCCAGAAATTTTC 1860
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 1861 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTATGCTGCTCCTCAGGTG 1920
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 1921 CAGCTCGAAAGAAATATCTTACTGTCTGTTTATGATATATGCTGCTCCTCAGGTGCTAGT 1980
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB 1981 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATATACCTAGCTCTCTAGGT 2040
QY 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
DB 2041 TATGTGTTTGTAGTAGTAGACAAACAGCGGATCTCTGTCCAGCGGCTTAAATTTGAAGC 2100
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
DB 2101 GCCTTTAAATATAAATGGGTCAATATAGAAATGACGATCAGGTGGAAGGACTCCAAATAT 2160
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740

Db 2161 CTAGCTTCGATATGATTTTCATTGACTTGTGGGATCCACGGCTGGTCTAT 2220
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2221 GGAGGATACCTCTCCCTGATGCAATTAATGCAGAGGTACAGATATCTTCAGGGTTCCTATT 2280
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2340
Qy 781 GlyHisProAspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2341 GGTCAACCCCTGACAGAATAACACGGGCTATTACTTAGGATCTGTGGCCATCGCAAGCAGAA 2400
Qy 801 LysPheProSerGlnProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2401 AAGTTCCTCTGACCAAAATCGTTTACTGTCTTACATGTTTCTCGATGAGATGTC 2460
Qy 821 HisPheAlaHisThrSerIleLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2461 CATTTTGCATACACCATATATTACTGAGTTTTTGTAGGGGTGGAAGCCATATGAT 2520
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2521 TTACAGATCTATCTCTCAGAGAGACACAGCATTAAGAGTTCCTGAATCTGGAGAACATTAT 2580
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2581 GAACTGCATCTTTTGTCACTACCTTCAAGAAAACCTTGATGATCATGTTGCTGCTTAAAA 2640
Qy 881 ValIle 882
Db 2641 GTGATA 2646

RESULT 2

US-10-170-789-39
; Sequence 39, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-170-789-39
Alignment Scores:
Pred. No.: 0 Length: 2649
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-825-632-1 (1-882) x US-10-170-789-39 (1-2649)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 1 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 60
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 61 GAGGAGAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTGTGACGGTAT 120
Qy 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 121 TCTGTGAGTCAGCTTTAAAAAGCTGTTCGCCGATACCAAGAAATATCATGCTACATATG 180
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 181 GCTAAGGCACCATCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 240
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 241 GACAGAATCTATTACCTTCCCATGTCTGTGAGAACACAGAGAAATATACACTGTTTTATTCT 300
Qy 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTyrTrpLysProLeuLeu 120
Db 301 GAAATTCCTCCAAACTATCAATAGCAGCAGCTTTAATGCTCTCTTGAAGCCCTTTTTC 360
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluLeuLeuGlySerArgIleAlaLeuLeuLys 880
 Db 2581 GAACTGTCATCTTTTGCACTACCTTCAAGAAAAACCTTGAGTACCGTATTTGCTGCTCTAAAA 2640
 Qy 881 ValIle 882
 Db 2641 GTGATA 2646

RESULT 3

US-09-976-674-2
 ; Sequence 2, Application US/09976674
 ; Patent No. US20020115843A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Qi, Steve
 ; APPLICANT: Akineanya, Karen
 ; APPLICANT: Riviere, Pierre
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
 ; FILE REFERENCE: 70669
 ; CURRENT APPLICATION NUMBER: US/09/976,674
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: US 60/240,117
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 2671
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-976-674-2

Alignment Scores:

Pred. No.: 0 Length: 2671
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0

US-10-825-632-1 (1-882) x US-09-976-674-2 (1-2671)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 Db 8 ATGGCAGCAATATGAAACAGACAGCTGGGTGGTGGATTTGAGATATTGAACTCGGACTGT 67
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 68 GAGGAGAATATTGAATCACAGGATCGCCCTAAATTTGGAGCCCTTTTATGTTGACGGGTAT 127
 Qy 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 128 TCCTGGAGTCAGCTTAAAGAGCTGTTCGCGATACAGAAATATCATGGCTACATGATG 187
 Qy 61 AlalysAlaProHisaspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 188 GCTAAGCACCACCATGATTTTCATGTTGTGAGAGAGATGATCCAGATGGACCTCATTTCA 247
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 248 GACAGAATCTATTACCTTGCCATGTCGTGGAGAACAGAGAAATATACACTGTTTATTTCT 307
 Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 308 GAAATTCCTCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGAAGCCCTCTTTTG 367
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 Db 368 GATCTTTTTCAGGCACACACTGGACTATGGAATGATTTCTCCGAGAGAGAACTATTAGA 427
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 428 GAAAGAAAACGCATTGGAAACAGTCGGAATTTGCTTCTTACGATTATCACCAAGGAAGTGA 487

Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 488 ACATTTCTGTTCACGCCGGTAGTGAATTTATACGTAAAGATGAGGGGCCACAAGGA 547
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 548 TTTACGCACACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTGCCACATACGATG 607
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 608 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTGCTTTATATACATAGCAACGATATT 667
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 Db 668 TGGATATCTAAACATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 727
 Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 728 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCAAGAAGAA 787
 Qy 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
 Db 788 TTTGATAGATATTCTGCTATTGCTGCTGCTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 847
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 Db 848 AAAATTCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 907
 Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 908 ACATCCCTATGTTGAAAACAGGAGGCGAGATTCAITCCGTTATCTCTTAAACACAGGTACA 967
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 968 GCAATCTTAAGTCACTTTTAAGATGTCAGAAAATATGATGCTGCTGAAGGAGGATC 1027
 Qy 341 IleAspValIleAspLysGluLeuIleGlnPropheGluIleLeuPheGluGlyValGlu 360
 Db 1028 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTTTGAAGAGGATTGAA 1087
 Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1088 TATATGCCAGACTGGATGGACTCTGAGGAAAATATGCTGGTGCATCTCTACTAGAT 1147
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1148 CGTCCCAAGACTCGCTCGCAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1207
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1208 GATGATGTTATGAAAAGGCAGAGACTCAITGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1267
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1268 ATTATCTATCAAGAAAACACAGACATCTCGATAAATATCCATGACATCTTTTCATGTGTTT 1327
 Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1328 CCCCAAGAGTCAGGAGGAAATTTGATTTATTTTGGCTCTGAATGCAAAAACAGGTTTC 1387
 Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1388 CGTCATTTATACAAATATCATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGGT 1447
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
 Db 1448 GGGCTGCTGCTCTCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGAGATAGCAATACCAGT 1507
 Qy 501 GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 Db 1508 GGTGAATGGGAAGTTCTTGGCCGGCATGATCTAATATCCAAGTTGATGAAGTCAGAAG 1567

521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
1568 CTGGTATATTTTGAAGGACCAAGAGCTCCCTTTAGAGCATACCTGTACGTAGTCAGT 1627
541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
1628 TACGTAAATCTCGAGAGGTGACACAGGCTGACTGACCGTGGCTACTACATCTTGTGTCG 1687
561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
1688 ATCAGTCAGCATCTGACTCTCTTTATTAAGTAAGTATAGTAACACAGAAATCCACACTGT 1747
581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
1748 GTGTCCTTTTCAAGACTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGAAATTT 1807
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPhe 620
1808 TGGGCCACATTTTGGATTGACAGGTCTCTCTGACTACTACTCTCCAGAAATTTTC 1867
621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
1868 TCTTTTGAAGTACTACTGGATTTTACATTTGATATGGATGCTCTACAGGCTCATGATCTA 1927
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGluValGlnLeu 660
1928 CAGCCTGGAAGAAATATCTTACTGTCTGTCTATATATGTTGGTCTCTCAGGTGCGATTG 1987
661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
1988 GTGAATAATTCGATTTAAAGAGAGTCAAGTATTTCCGCTTGAATACCTACGCTCTCAGGT 2047
681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
2048 TATGTGTTTGTAGTGATAGACACACAGGGGATCTGTCTCACCAGGGGCTTAAATTTGAAGGC 2107
701 AlaPheLysTyrLysMetGlyGlnIleGluLysLeuAspGlnValGluGlyLeuGlnTyr 720
2108 GCCTTTAATATAAATGGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCCAATAT 2167
721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyr 740
2168 CTAGCTTCTCGATATGATTTTCACTTACTGACTCTAGATCGTGTGGGCATCCACGGCTGCTAT 2227
741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaLys 760
2228 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTACAGATATCTTCAGGGTTGCTATT 2287
761 AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
2288 GCTGGGGCCCCAGTCACTCTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2347
781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
2348 GGTACCCCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2407
801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
2408 AAGTTCCTCTGCAACCAATCGTTTACTGCTCTTACATGTTTTCCTGGATGAGAATGTC 2467
821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
2468 CATTTTGCATATACCATGATATTTACTGAGTTTTTTAGTGGGGCTGGAAGCCATATGAT 2527
841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
2528 TTACAGATCTATCTCAGAGAGACACAGCATAGAGTTCTCGAATCGGGAGACATATAT 2587
861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
2588 GAATGTCATCTTTTGGCACTACCTTCAAGAAACCTTGGATCACGTTATGCTGCTCTAAAA 2647
881 ValIle 882

Db 2648 GTGATA 2653
RESULT 4
US-10-982-512-2
; Sequence 2, Application US/10982512
; Publication No. US2005009081A1
; GENERAL INFORMATION:
; APPLICANT: Q1, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Jean-Louis
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982, 512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976, 674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240, 117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-2
Alignment Scores:
Prod. No.: 0 Length: 2671
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0
US-10-825-632-1 (1-882) x US-10-982-512-2 (1-2671)
Qy 1 MetaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 8 ATGGCAGCAGCATATGGAACACAGACAGCTGGGTGTGAGATATTTGAACTCGGACTGT 67
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 68 GAGGGAATATTGAATACACAGATCGGCTTAATTTGAGGCTTTTATGTTGAGCGGTAT 127
Qy 41 SerTyrSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 128 TCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAAGAAATATCATGCTACATGATG 187
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 188 GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATCA 247
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAenThrLeuPheTyrSer 100
Db 248 GACAGAAATCTATTACCTTGCCATGTCTGTGAGAACACAGAGAAATACACTGTTTATCT 307
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
Db 308 GAATTTCCCAAAATATCAATAGACAGCAGCTTAATGCTCTCTTGGAGCCCTCTTTTG 367
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 368 GATCTTTTTCAGGCAACACTGGACTATGGAATGTTATCTCGAGAAAGAACTATTAAGA 427
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 428 GAAGAAGAACCATTCGACAGCTCGGAATGCTCTTACGATTATCACCAAGGAAGTGA 487
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 488 ACATTTCTGTTTTCAGCCGGTAGTGGATTTTATCACGTAAAGATGGAGGCCCAAGGA 547

Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 548 TTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATG 607
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 608 GATCCAAAATTATGCGCTGCTGATCCAGACTGGATGCTTTATACATAGCAACGATATT 667
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTrpValHisAsnGluLeu 240
Db 668 TGGATATCTPAACATCGTAACGAGACGAGAGACTCACTTATGTGCACATGAGCTA 727
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 728 GCCAACATCGAAGAAGATGCCAGATCAGCTGCGAGTCGCTACCTTGTCTCCAAAGAA 787
Qy 261 PheAspArgTrpSerGlyTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 788 TTTGATAGATATTCTGGCTATTGGTGTGTTGCCAAAGCTGAAACAACTCCCACTGGTGT 847
Qy 281 LysIleLeuArgIleLeuTrpGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 848 AAAATCTTAGAATTTCTATAGAGAAATGATGAATCTGAGGTGGAAATATTTCATGTT 907
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTrpProLysThrGlyThr 320
Db 908 ACATCCCTTATGTTGGAAACAGAGGGCGAGATTCATTCCTTATCCCTAAACAGGTACA 967
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 968 GCAATCTCTAAAGTCACCTTTAAGATGTGAGAAATATGATGATGCTGAGGAGGATC 1027
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluValGlu 360
Db 1028 ATAGATGTCATAGAAAGCAATAATCAACCTTTGAGATTCATTGTAAGGAGTTGAA 1087
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIleLeuLeuAsp 380
Db 1088 TATATTGCCAGACTGGATGGACTCTCGAGGAAATATGCTGGTCCATCTCTAGAT 1147
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1148 CGCTCCAGACTCGCTCGCAGATAGTGTGATCTCACCTGAATTAATTAATCCAGTAGAA 1207
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1208 GATGATGTTATGGAAGGCAGAGACTCAATTGAGTCAGTCGCTGATTCGTGACGCCACTA 1267
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1268 ATTATCTATGAGAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1327
Qy 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1328 CCCCAAGTCACGAGAGGAATTCAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1387
Qy 461 ArgHisLeuTrpLysIleThrSerIleLeuLysGluSerLysTrpLysArgSerSerGly 480
Db 1388 CGTCAVTTATACAAAATTACATCTATTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT 1447
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 1448 GGGCTCGCTGCTCCAAAGTATTTCAAGTGTCTATCAAAAGGAGATAGCAATTAACAGT 1507
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1508 GGTGAATGGAGTTCTTGGCCGGCATGGATCTATATCCAAAGTTGATGAATCAGAGG 1567
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTrpValValSer 540
Db 1568 CTGATATATTGGAAGGCCCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGT 1627

RESULT 5

Qy 541 TyrValAsnProGlyGluValThrArgIleThrAspArgGlyTyrSerHisSerCysCys 560
Db 1628 TACGTAAATCCCTGGAGAGGTGACAAGGCTGACTACCGTGGCTACTCACATTTCTTGTCTG 1687
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTrpSerAsnGlnLysAsnProHisCys 580
Db 1688 ATCAGTCACGACTGTGACTTCTTTATAGTAAGTATAGTAACACGAGAAGATCCACACTGT 1747
Qy 581 ValSerLeuTrpLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1748 GTGTCCCTTTACAAGCTATCAAGTCTCAAGATCACCAACTGTGCAAAACAAAGGAATTT 1807
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTrpThrProProGluIlePhe 620
Db 1808 TGGGCCACCATTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC 1867
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTrpGlyMetLeuTrpLysProHisAspLeu 640
Db 1868 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAAGCCTCATGATCTA 1927
Qy 641 GlnProGlyLysLysTrpProThrValLeuPheIleTrpGlyGlyProGlnValGlnLeu 660
Db 1928 CAGCCTGGAAAGAAATATCTACTGTCTGTTCATATATGTTGCTCTCAGGTGCTAGTTG 1987
Qy 661 ValAsnAsnArgPheLysGlyValLysTrpPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1988 GTGAATAATCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGT 2047
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2048 TATGTGGTTGTAGTAGACAAACAGGGGATCTGTCCACCGAGGCTTAAATTTGAAGGC 2107
Qy 701 AlaPheLysTrpLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTrp 720
Db 2108 GCCTTTAATATAAATGSGTCAANTAGAAATTCAGATCAGGTGGAAGGACTCCAATAT 2167
Qy 721 LeuAlaSerArgTrpAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTrp 740
Db 2168 CTAGCTTCTCGATATGATTTTCACTTAGATCTGTGGGATCCACGGCTGTCTCTAT 2227
Qy 741 GlyGlyTrpLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2228 GGAAGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATT 2287
Qy 761 AlaGlyValaProValThrLeuTrpIlePheTrpAspThrGlyTrpThrGluArgTrpMet 780
Db 2288 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2347
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTrpTrpLeuGlySerValAlaMetGlnAlaGlu 800
Db 2348 GGTCACTGACCCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGCAGAA 2407
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2408 AAGTTCCTCTGAAACAAATCGTTTACTGTCTTACATGGTTTCTCTGGATGAGAATGTC 2467
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTrpAsp 840
Db 2468 CATTTTGCACATACCAAGTATATTAATGAGTTTGTAGGGCTGGAAACCATATGAT 2527
Qy 841 LeuGlnIleTrpProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTrp 860
Db 2528 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTAT 2587
Qy 861 GluLeuHisLeuLeuHisTrpTrpLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2588 GAACGTGATCTTTTGTGACTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAA 2647
Qy 881 ValIle 882
Db 2648 GTGATA 2653

US-10-311-035-30		US-10-825-632-1 (1-982) x US-10-311-035-30 (1-3106)	
; Sequence 30, Application US/10311035		Qy 1 MetAlaAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100	
; Publication No. US20040023243A1		Db 443 GACAGATCTATTACCTTGCATGTCTGGTGAGACAGAGAAATACACTGTTTTATTCT 502	
; GENERAL INFORMATION:		Qy 101 GluIleProLysThrIleAsnArgAlaAlaValMetLeuSerTrpLysProLeuLeu 120	
; APPLICANT: INCYTE GENOMICS, INC.		Db 503 GAAATCCCAAAATATCAATAGAGCAGCAGTCTTAATGCTCTCTTTGMAAGCTCTTTTG 562	
; APPLICANT: YUE, Henry		Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140	
; APPLICANT: ELLIOTT, Vicki		Db 563 GATCTTTTTCAGGCAACACTGGACTATGGAAATTCCTTACGATTATCACAAGAGACTATTAAAG 622	
; APPLICANT: GANDHI, Ameena R.		Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160	
; APPLICANT: LAL, Preeti		Db 623 GAAAGAAAACGATTTGGAACTCGGAATTCCTTACGATTATCACAAGAGACTATTAAAG 682	
; APPLICANT: AU-YOUNG, Janice		Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180	
; APPLICANT: TRIBOULEY, Catherine M.		Db 683 ACATTTCTGTTCGAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAAAGGA 742	
; APPLICANT: DELEGEANE, Angelo M.		Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200	
; APPLICANT: BAUGHN, Mariah R.		Db 743 TTTTACGCAACCAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCTCCCAACATACGAGTG 802	
; APPLICANT: NGUYEN, Dannel B.		Qy 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220	
; APPLICANT: LEE, Ernestine A.		Db 803 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 862	
; APPLICANT: HAFALIA, April		Qy 221 TrrPheSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240	
; APPLICANT: KHAN, Farrah A.		Db 863 TGGATATCTTAACATCGTAACAGAGAAAGAGAGACTCATTATGTGCAATGAGCTA 922	
; APPLICANT: CHAWLA, Narinder K.		Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260	
; APPLICANT: YAO, Monique G.		Db 923 GCCACATCGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCCAAGAGAA 982	
; APPLICANT: LU, Dyung Aina M.		Qy 261 PheAspArgTyrSerGlyTyrTrrPtyCysProLysAlaGluThrThrProSerGlyGly 280	
; APPLICANT: ARVIZU, Chandra S.		Db 983 TTTGATAGATATTTCTGGCTATTTGGTGTTGTCAAAAGCTGAAAACACTCCCACTGGTGGT 1042	
; APPLICANT: TANG, Y. Tom		Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300	
; APPLICANT: WALSH, Roderick T.		Db 1043 AAAATCTTTAGAAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTTCATGTT 1102	
; APPLICANT: AZINZAL, Yalda		Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320	
; APPLICANT: Lu, Yan		Db 1103 ACATCCCTTATGTTGGAACCAAGGAGGCGCAGATTTCATTCCGTTATCTCTAAACAGGTACA 1162	
; APPLICANT: RAMKOMAR, Jayalaximi		Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340	
; APPLICANT: REDDY, Roopa		Db 1163 GCAATCTCTAAAGTCACTTTTAAAGATGTGAGAAATAATGATGATGCTGAAGGAAGGATC 1222	
; APPLICANT: DAS, Depopriya		Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360	
; APPLICANT: KEARNEY, Liam		Db 1223 ATAGATGTCATAGATAAGGAACTAATTCACCTTTTGAGATTCATTATTTGAAGGAGTTGAA 1282	
; APPLICANT: KALLICK, Deborah A.		Qy 361 TyrIleAlaArgAlaGlyTrrThrProGluGlyLysTyrAlaTrrSerIleLeuLeuAsp 380	
; TITLE OF INVENTION: Proteases		Db 1283 TATATTTGCCAGAGCTGGATGGACTCTCGAGGGAATAATGCTTGGTCCATCTACTAGAT 1342	
; FILE REFERENCE: PI-0123 PCT		Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400	
; CURRENT APPLICATION NUMBER: US/10/311.035		Db 1343 CGCTCCGAGCTCGCTACAGATAGTGTGTGATCTCACCTGATTTATTTATCCCAAGTAGAA 1402	
; CURRENT FILING DATE: 2002-12-10		Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420	
; PRIOR APPLICATION NUMBER: 60/213,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946		Db 1403 GATGATGTTATGGAAGGCGAGACTCAITGAGTCAGTGGCTGATTTCTGTGAGGCCACTA 1462	
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14		Qy 421 IleIleTyrGluGluThrThrAspIleTrrPheAsnIleHisAspIlePheHisValPhe 440	
; NUMBER OF SEQ ID NOS: 42		Db 1463 ATTATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1522	
; SOFTWARE: PERL Program			
; SEQ ID NO 30			
; LENGTH: 3106			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CB1			
US-10-311-035-30			
Alignment Scores:			
Pred. No.:	0	Length:	3106
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0
US-10-825-632-1 (1-982) x US-10-311-035-30 (1-3106)			
Qy	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	203	ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTGAACCTGGGACTGT	262
Qy	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Db	263	GAGGAGAAATATTGAATTCACAGGATCGGCTAAATTTGGAGGCTTTTATGTTGAGCCGGTAT	322
Qy	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Db	323	TCCTGGAGTCAGCTTAAAGGCTGCTTCGCGATACCAAGAAATATCATGCTACATGATG	382
Qy	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Db	383	GCTTAAGGCCACCATGATTTTCATGTTTGTGTAAGGAAGTAATGATCCAGATGACCTCATTC	442

QY 441 ProGlnSerHisGluGluGluIleGluPheAlaSerGluCysGlyThrGlyPhe 460
Db 1523 CCCCAAGTCACGAAGAGGAAATTGAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTC 1582
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1583 CGTCATTATATACAAATTTATCATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGT 1642
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1643 GGGCTGCTGCTCCAGTGATTTCAAGTGCTCTCAAGAGAGAGATAGCAATTTACCACT 1702
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
Db 1703 GGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATATCAAGTTGATGAAGTCAGAAGG 1762
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1763 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCCTGTACGTAGTCAGT 1822
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1823 TACGTAATCTCGAGAGGGTGACAAGGCTGACTGACCGTGGCTACTCACTTTCTTCCTGCG 1882
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1883 ATCAGTCAGCACTGTGACTTCTTTATTAAGTAAAGTATAGTAAACCAAGAGAAATCCACTGT 1942
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1943 GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAAACAAAGGAATTT 2002
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2003 TGGGCCACCATTTTGGATTACAGCAGGTCCTTCTCCGACTACTACTCTCCAGAAATTTTC 2062
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2063 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGAGTGTCTTACAGGCTCATGATCTA 2122
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2123 CAGCTCGAAGAAATATCTTACTGTCTGTTCATATATGCTGCTCCTCAGGTGCAAGTTG 2182
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2183 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCGGCTTGAATACCCCTAGCCCTCTAGGT 2242
QY 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLysLysPheGluGly 700
Db 2243 TATGTGGTTGTAGTATAGACAACAGGGGATCCCTGTCCAGGGGCTTAAATTTGAGGC 2302
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyr 720
Db 2303 GCCTTTAAATATAAATGGGTCAATAGAAATGACCATCAGGTGGAAGGACTCCCAATAT 2362
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2363 CTAGCTTCTCGATATGATTTTCATTGACTAGATCGTGTGGGCATCCACGGCTGCTCTAT 2422
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2423 GGAGGATACCTCTCCCTGATGCAATTAATGACAGAGGTTCAGATATCTTCAGGGTGTCTATT 2482
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2483 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGACGTTATATG 2542
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2543 GGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2602
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820

Db 2603 AAGTTCCCTCTGAACCAATTCGTTTACTGCTCTTACATGGTTTCTCGATGAATGTC 2662
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2663 CATTTTGCACATACCACTATATTTACTGAGTTTTTTAGTGAGGGCTGAAAGCCCATATGAT 2722
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2723 TTACAGATCTATCTCTCAGGAGAGACAGACATAGAGTTTCTGAAATCGGGAGAACATTAT 2782
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2783 GAACGTGCATCTTTTGCACTACTCTTCAAGAAAACTTGGATCAGTATTGCTGCTCTAAAA 2842
QY 881 ValIle 882
Db 2843 GTGATA 2848

RESULT 6
US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FP15217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-415-122-5

Alignment Scores:
Pred. No.: 0 Length: 3120
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-825-632-1 (1-882) x US-10-415-122-5 (1-3120)

QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATCGAAACACAGACAGCTGGGTGTTGAGATATTTGAAACTGCGGACTGT 273
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATTTGAATTCACAGGATCGGCTTAAATTTGAGCCCTTTTATGTTGAGCGGTAT 333
QY 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCTTGAGTCACTTAAAAAGCTGCTCCGATACCAAGAAATATCATGCTACATCATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCACCATCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAanThrLeuPheTyrSer 100
Db 454 GACAGAACTTATTTACCTGCTCCATGTTGGTGGAGAACAGAGAAATATACACTGTTTATCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCCAAACTATCATAGACAGCAGTCTTAATGCTCTCTTGAAGCCCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

Db 574 GATCTTTTTCAGGCCAACCTGGACTATGGAAATGCTATTCTCTCGAAGAGAACTATTAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAGAAGAACCATTTGAACAGCTGGAAATTCCTCTTACGATATATACACAGGAAGTGGG 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTCGAAGCCGCTAGTGGAAATTTATCAGCTAAAGATGGAGGGCCCAAGGA 753
Qy 181 PheThrGlnGlnProLeuAArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACCTAGTTGTCCCAACATACGGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACAGATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAACATCGTAACAGAGAAGAGAGACTCACATTATGTGCACAAATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 934 GCCAACATGGGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAAAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATCTGGCTATTTGGTGTGTCCAAAGCTGAAACCACTCCCAAGTGTGGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 1054 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAATATATCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCCATGTTTGAAGAACAGAGGGCAGATTCAATCCGTATCTCTTAAACACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAATCTCTAAAGTCACCTTTTAAGATGTCCAGAAATATGATGATGCTGGAAGGAGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCTAGATAAGGAACCTAAATCAACCTTTTGGAGATCTATTTGAAGAGTTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGAGCTGGATGACTCTCTGAGGGAAAAATATGCTTGGTCCATCTCTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACTGAAATTTATATCCAGTAGAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAAGGCAGAGCTCATTTGAGTCAGTGCCTGATTCGTGACGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1534 CCCCAGAGTCACGAAGAGGAAATTTAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 GGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGCTGCCTGCTCCCAAGTGATTTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAGT 1713

Qy 501 GlyGluTyrProGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTTCTTTGGCCGCATGGATCTAATATCCAGTTGATGAGTCAGAAG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGCCCAAGACTCCCTTTTAGAGCATCCTGTACGTAGTCAAT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TACGTAAATCTCTGAGAGAGTGACAAGGCTGACGCGGTGGCTACTCACATTTCTGCTGC 1893
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAAGCATGTGACTCTTTTATTAAGTAAGTATAGTAACCAAGAAATCCACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCCTTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCAGAAACAAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPhe 620
Db 2014 TGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTTCTTGACTATATCTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTTACTGGATTTTACATTTGATGGGATGCTCTACAAGCTCATGATCTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCCTGGAAAGAAATATCTTACTGTCTGTTCATATATGTTGGTCTCTCAGGTGCAAGTTG 2193
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2194 GTGAATATCCGTTTAAAGGAGTCAGTATTTCCGTTGAAATACCTTAGCTCTCTAGGT 2253
Qy 681 TyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
Db 2254 TATGTGGTGTAGTGATAGACAACAGGGGATCTGTCCAGAGGGCTTAAATTTGAAGGC 2313
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2314 GCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT 2373
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2374 CTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTAT 2433
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2434 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCCAGATATCTTCAGGGTGTCTATT 2493
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2554 GGTCACTCTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2614 AAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGCTTTCTTGATGAGAATGC 2673
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2674 CAATTTGACATACACAGTATATTTACTGAGTTTTTTTAGTGGGCTGGAAGCCATATGAT 2733
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2734 TTTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTTCTCTGAATCGGAGAACATTAT 2793

Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
 Db 2794 GAACTGCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGTATTCGTCTTAAAA 2853
 Qy 881 ValIle 882
 Db 2854 GTGATA 2859

RESULT 7

US-10-825-632-2
 ; Sequence 2, Application US/10825632.
 ; Publication No. US20040191826A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ABBOTT, Catherine Anne
 ; APPLICANT: GORRELL, Mark Douglas
 ; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
 ; FILE REFERENCE: PCSB-100-Div. 1
 ; CURRENT APPLICATION NUMBER: US/10/825,632
 ; CURRENT FILING DATE: 2004-04-15
 ; PRIOR APPLICATION NUMBER: US 10/070,464
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: PCT/AU00/01085
 ; PRIOR FILING DATE: 2000-09-11
 ; PRIOR APPLICATION NUMBER: AU PQ5709
 ; PRIOR FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: AU PQ2762
 ; PRIOR FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 3120
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 US-10-825-632-2

Alignment Scores:
 Pred. No.: 0 Length: 3120
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-825-632-1 (1-882) x US-10-825-632-2 (1-3120)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 Db 214 ATGCGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGAACTGCGGACTGT 273
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 274 GAGAGATATTGGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
 Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 334 TCCTGGAGTCAGCTTAAAGAGCTGTTCGCGATACAGAAATATCATCGCTACATCATG 393
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 454 GACAGATCTATTACCTTGGCATGCTGGTGAGAACAGAGAAATACATGTTTTATTCT 513
 Qy 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 514 GAAATCCCAAACTATCATAGAGCAGCAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG 573
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
 Db 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTTCGAGAGAGAACTATTAAAGA 633

Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 634 GAAAGAAAACGCAATTTGGAAACAGTCGGAATCTCTTACGATTATACCAAGGAAGTGA 693
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 694 ACATTTCTGTTTCNAGCCGCTAGTGTGAAATTTATCACGTAAGAGATGAGGCGCCACA 753
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 754 TTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGATG 813
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 814 GATCCAAAATATATGCCCGCTGATCCAGATGGATTGCTTTTATACATACAGACATATT 873
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 Db 874 TGGATATCTAACATCGTAACACAGAGAGAGAGAGACTCACTTATGTGCACATGAGCTA 933
 Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValIleAlaThrPheValLeuGlnGlu 260
 Db 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAA 993
 Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGly 280
 Db 994 TTTGATAGATATTTCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT 1053
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
 Db 1054 AAAATTTCTAGAAATCTATATGAGAAATGATGAACTGAGGTGGAAATATTATCATGTT 1113
 Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 1114 ACATCCCTATGTTGGAAACAAGAGGCGCAGATTCACTCCGTATCTCTAAACAGAGTACA 1173
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 1174 GCAAAATCTTAAAGTCACTTTTAAAGATGTGAGAAATATGATGATGCTGAGGAGAGATC 1233
 Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1234 ATAGATGTCATAGATAGGAACTAATTCACCTTTTGAGATCTATTGAGGAGTTGAA 1293
 Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1294 TATATTGCCAGAGCTGGATGGACTCTCGAGGAAATATGCTGGTCCATCTACTAGAT 1353
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuLeuSerProGluLeuPheIleProValGlu 400
 Db 1354 CGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTTATCCAGTAGAA 1413
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1414 GATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1474 ATTATCTATGAGAAACAACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
 Qy 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1534 CCCCAAGTCCAGAGAGAAATTTGAGTTATTTTTCCTCTGCTGTAATGCAAAACAGGTTC 1593
 Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGT 1653
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
 Db 1654 GGGCTGCTGCTCCAGTGTATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAGT 1713
 Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520

Db	1714	GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG	1773	
Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHieLysTyrValValSer	540	
Db	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTAGTAGTCAGT	1833	
Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	
Db	1834	TACGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACATCTTGTCTGC	1893	
Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	
Db	1894	ATCAGTCAGCACTGTGACTTCTTTAAGTAGTATAGTAAACAGAGNAATCCACACTGT	1953	
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	
Db	1954	GTGTCCCTTTTCAAGACTATCAAGTCCTGAAGATGACCACTTGCAAAACAAGNAATTT	2013	
Qy	601	TtpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	
Db	2014	TGGGCCACCATTTTGGATTACAGCAGGTCTCTTCCGTACTACTCTCCAGAAATTTTC	2073	
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	
Db	2074	TCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTACAGCCTCATGATCTA	2133	
Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660	
Db	2134	CAGCCTGGAAGAAATATCTTACTGTCTGTTCATATATGTTGGTGGTCTCAGGTGCAGTTG	2193	
Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680	
Db	2194	GTGAATATTCGGTTTAAAGGAGTCAAGTATTTCCGGTTGAATACCCCTAGCCTCTTAGGT	2253	
Qy	681	TyrValValValValIleAsnAsnArgLysSerCysHisArgGlyLeuLysPheGluGly	700	
Db	2254	TATGTGTTTGTAGTATAGACACACAGGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGC	2313	
Qy	701	AlaPheLysTyrLysMetGlyGlnIleLeuLeuAspGlnValGluGlyLeuGlnTyr	720	
Db	2314	GCCTTTTAAATATAAATGGTCAATAGAAATTCAGATCAGTGGAGGACTCCCAATAT	2373	
Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyLysGlyTyrSerTyr	740	
Db	2374	CTAGCTTCTCGATATGATTTTCACTTGTAGTCTGTGGGCATCCACGGCTGGTCCAT	2433	
Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760	
Db	2434	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAAGATATCTTCAGGGTGGTATT	2493	
Qy	761	AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780	
Db	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2553	
Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800	
Db	2554	GGTCACTCCCTGACCAGAAATCAACAGGCTATTACTTAGGATCTGTGGCCATCGAAGCAGAA	2613	
Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820	
Db	2614	AAGTTCCTCTCAACCAATCGTTTACTGCTCTTACATGCTGTTTCTCGATGAGAATGTC	2673	
Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840	
Db	2674	CATTTTGCATACACAGTATATTTACTAGTTTTTTGTAGGGCTGGAAAGCCCATGAT	2733	
Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860	
Db	2734	TTACAGATCTATCTCAGGAGACACACAGTAAAGTTCCTGAATCGGAGAACATTAT	2793	
Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnLeuAsnLeuGlySerArgIleAlaLeuLys	880	

Db2794GAACTGCATCTTTTGCACTACTCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA2853

Qy881ValIle882

Db2854GTGATA2859

RESULT 8

US-10-170-789-37

; Sequence 37, Application US/10170789

; Publication No. US20030180930A1

; GENERAL INFORMATION:

; APPLICANT: Rachel E. Meyers

; APPLICANT: Olandt, Peter J.

; APPLICANT: Kapeller-Libermann, Rosana

; APPLICANT: Curtis, Rory A. J.

; APPLICANT: Williamson, Mark

; APPLICANT: Welch, Nadine

; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,

; AND PROTEASE FAMILY MEMBERS AND USES THEREOF

; FILE REFERENCE: 10448-191001

; CURRENT APPLICATION NUMBER: US/10/170,789

; CURRENT FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: US 09/797,039

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: PCT/US01/06525

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 60/186,061

; PRIOR FILING DATE: 2000-02-29

; PRIOR APPLICATION NUMBER: US 09/882,166

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: PCT/US01/19269

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 60/212,078

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: US 09/934,406

; PRIOR FILING DATE: 2001-08-21

; PRIOR APPLICATION NUMBER: PCT/US01/26052

; PRIOR FILING DATE: 2001-08-21

; PRIOR APPLICATION NUMBER: US 60/226,740

; PRIOR FILING DATE: 2000-08-21

; PRIOR APPLICATION NUMBER: US 09/861,801

; PRIOR FILING DATE: 2001-05-21

; PRIOR APPLICATION NUMBER: PCT/US01/16549

; PRIOR FILING DATE: 2001-05-21

; PRIOR APPLICATION NUMBER: US 60/205,508

; PRIOR FILING DATE: 2000-05-19

; PRIOR APPLICATION NUMBER: US 09/801,267

; PRIOR FILING DATE: 2001-03-06

; PRIOR APPLICATION NUMBER: PCT/US01/07138

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: US 60/187,454

; PRIOR FILING DATE: 2000-03-07

; PRIOR APPLICATION NUMBER: US 09/829,671

; PRIOR FILING DATE: 2001-04-10

; PRIOR APPLICATION NUMBER: PCT/US01/40483

; PRIOR FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: US 60/197,508

; PRIOR FILING DATE: 2000-04-18

; PRIOR APPLICATION NUMBER: US 09/961,721

; PRIOR FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: PCT/US01/29904

; PRIOR FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: US 60/235,023

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US 10/045,367

; PRIOR FILING DATE: 2001-11-07

; PRIOR APPLICATION NUMBER: US 60/246,561

; PRIOR FILING DATE: 2000-11-07

; PRIOR APPLICATION NUMBER: US 09/801,275

; PRIOR FILING DATE: 2001-03-06

; PRIOR APPLICATION NUMBER: PCT/US01/07074

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: US 60/187,420

; PRIOR FILING DATE: 2000-03-07
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: Fast-Seq for Windows Version 4.0
 ; SEQ ID NO 37

; LENGTH: 3143
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (229) ... (2874)
 US-10-170-789-37

Alignment Scores:

Pred. No.: 0 Length: 3143
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-825-632-1 (1-882) x US-10-170-789-37 (1-3143)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 Db 229 ATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTGAAACCTGGCGACTGT 288
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 289 GAGGAGNATATTGAAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATGTGACGGTAT 348
 Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 349 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACCAAGAAATATCATGGCTACATGATG 408
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 409 GCTAAGGCACCATGATTTTCATGTTTGTGTAAGAGAGAAATGATCCAGATGGACCTCATTTCA 468
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 469 GACAGAAATCATTAACCTTGCATGCTGTGGTGAACAGAGAAATATACACTGTTTATCT 528
 Qy 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 529 GAATATCCCAAACTATCATAGAGCAGCAGCTTAATGCTCTCTTGGAGCCCTCTTTT 588
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 Db 589 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAGA 648
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 649 GAAAGAAACAGCAATGGACAGTCGGAATTTGCTTCTTACGATTATCACCAAGGAAGTGA 708
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 709 ACATTTCTGTTTACGCCGGTAGTGGAATTTATCAGTAAAGATGGAGGCCACACAGGA 768
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 769 TTTACGCAACAACCTTTAAGGCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATG 828
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 829 GATCCAAAATATTGCCCTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 888
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeu 240
 Db 889 TGGATATCTTAACATCGTAACACAGAGAAGAAAGAGACTCATCTTATGTCCACATGAGCTA 948
 Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 949 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA 1008

Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 Db 1009 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACCACTCCCAAGTGGTGGT 1068
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
 Db 1069 AAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1128
 Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 1129 ACATCCCTATGTTGGAACAAGGAGGCGAGATTCAATCCGTTATCTCTAAACAGGTACA 1188
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 1189 GCAAAATCCTAAATCACTTTTAAGATGTGAGAAATAATGATGATGCTGTGCTCATCTACTAGAT 1248
 Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1249 ATAGATGTCATAGATAAGGAACCTAATTCAACCTTTTGAGATTTCTATTTGAAGGAGTTGAA 1308
 Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1309 TATATGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTGTGCTCATCTACTAGAT 1368
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1369 CGCTCCAGACTCGCTGCAGATAGTGTGATCTCACCTGAAATATTTATCCAGTAGAA 1428
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1429 GATGATGTTATGAAAGGAGGAGACTCATTTAGTCAGTGCCTGATTTCTGTGACGCCACTA 1488
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1489 ATTATCTATGGAAGAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1548
 Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1549 CCCCAGAGTCACAGAGGAAATGAGTTATTTTTCCTCTGAATGCAAAACAGGTTTC 1608
 Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1609 CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAATATAAAGCATCCAGTGGT 1668
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
 Db 1669 GGGCTGCTCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGAGATAGCAATACCAGT 1728
 Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 Db 1729 GGTGAATGGGAGTTCTTGCCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1788
 Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleSerTyrValValSer 540
 Db 1789 CTGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCACT 1848
 Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 Db 1849 TACGTAAATCTCGAGAGGTGCAAGCGCTGACTGACCGGTGGCTACTTCACATTTCTTGTCTGC 1908
 Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 Db 1909 ATCAGTCAGCACTGTGACTCTTTTAAAGTAGTATAGTAGTAAACAGAGAAATCCACACTGT 1968
 Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
 Db 1969 GTGTCCCTTTTACAGCTATCAAGTCTCTGAAGATGACCAACTTGAACCAACAAAGGAATTT 2028
 Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLeuIlePhe 620
 Db 2029 TGGGCCCACTTTTGGATTTCAGCAGGTCTCTTCTCTGATATATCTCTCTCCCAAAATTTTC 2088

QY	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrIleProHisAspLeu	640
Db	2089	TCCTTTGAAGTACTACTCGGATTTACATTTGATGGGATGCTCTCAAGACCTCATGATCTA	2148
QY	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu	660
Db	2149	CAGCCTGGAAGAAATATCTCTACTGTCTCTCATATATGGTGGTCTCCTCAGGTGAGTTG	2208
QY	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	2209	GTGAATAATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGT	2268
QY	681	TyrValValValIleAspAsnArgLysTyrCysHisArgGlyLeuLysPheGluGly	700
Db	2269	TATGTGGTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGGC	2328
QY	701	AlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyr	720
Db	2329	GCCTTTAAATATAAATGGGTCAANTAGAAATTCAGCATCAGGTGGAAGACTCCCAATAT	2388
QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
Db	2389	CTAGCTTCTCGATATGATTTCAITTGACTTTAGATCTGTGTGGCATCCACGGCTGGTCCCTAT	2448
QY	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2449	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAATATCTTCAGGGTTCGATTT	2508
QY	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2509	GCTGGGGCCCCAGTCACTCTCTGGATCTCTTATGATACAGGATACACGGAACTGATATG	2568
QY	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	2569	GGTCACCTTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2628
QY	801	LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2629	AGTTTCCCTCTGAAACCAATCGTTTACTGTCTTACATGGTTTCTCTGGATGAGATGTC	2688
QY	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2689	CATTTTGCATACACAGTATATTACTGAGTTTTTTAGTGGGCTGGAAGCCATATGAT	2748
QY	841	LeuGlnIleTyrProGlnArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2749	TTACGATCTATCTCGAGAGACACAGCATAGAGTTCTGTAAATCGGAGAACATTAT	2808
QY	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
Db	2809	GAATCGATCTTTTGCACTTACCCTTCAGAAACCTTGATCACGATTTGCTCTTAAA	2868
QY	881	ValIle 882	
Db	2869	GTGATA 2874	
RESULT 9			
US-09-976-674-12			
; Sequence 12, Application US/09976674			
; Patent No. US20020115943A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: Patentin version 3.1			
;			
; SEQ ID NO 12			
; LENGTH: 4829			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-976-674-12			
Alignment Scores:			
Pred. No.:	0	Length:	4829
Score:	4680.00	Matches:	882
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	0
Query Match:	99.6%	Indels:	2
DB:	3	Gaps:	0
US-10-825-632-1 (1-882) x US-09-976-674-12 (1-4829)			
QY	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	214	ATGGCAGCAGCAATGGAAACAGAACACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT	273
QY	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Db	274	GAGGAGAATATTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTGAGCGGTAT	333
QY	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Db	334	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACCAGAAAAATATCATGCTACATGATG	393
QY	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Db	394	GCTTAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTC	453
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Db	454	GACAGAACTTATTACTTGGCCATGCTGTGTGAGAACAGAGAAAAATACACTGTTTTATTCT	513
QY	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Db	514	GAAATTTCCAAACATCATCAATAGAGCAGCAGCTTAAATGCTCTCTTGGAAAGCCTCTTTG	573
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Db	574	GATCTTTTTTCAGGCAACACTCGACTATGGAATGTAATCTCGAAGAAGAACTATTAAAG	633
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Db	634	GAAAGAAAAACGCAATTTGGAACTCGGAATTCCTTACGATTATCACCAAGGAAGTGA	693
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Db	694	ACATTTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAGATGGAGGGCCACAGGA	753
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Db	754	TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGATG	813
QY	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Db	814	GATCCAAATATATGCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	873
QY	221	TrpIleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeu	240
Db	874	TGGATATCTAACATCTGTAAACAGAGAGAGGAGACTCACTTATGTGCACAATGAGCTA	933
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
Db	934	GCCAACTGGAAGAAGATCCAGATCAGCTGAGTGGCTACTTTGTTCTCCAAGAGAA	993
QY	261	PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly	280
Db	994	TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGTGT	1053
QY	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300

```
|||||
1054 AAAATCTTAGAATCTTATATGAAGAAATGATGAATCTGAGGTGGAATATTTCATGTT 1113
Db
301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY
1114 ACATCCCTCATGTTGGAAACACAGAGGGCAGATTCATCCGTTATCCCTAAACAGGTACA 1173
Db
321 AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle 340
QY
1174 GCAATCTCTAAAGTCATCTTTAGATGTGAGAAATANTGATGTGCTGAAGGAAGATC 1233
Db
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY
1234 ATAGATGTCTAGATAGAGAACTAAATCAACCTTTTGAGATCTTATTTGAAGAGGTTGAA 1293
Db
361 TyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY
1294 TATATTGCCAGAGCTGGATGGACTCCCTGAGGGGAAATATGCTTGGTGCATCTCTACTAGT 1353
Db
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY
1354 CGCTCCAGACTCCCTTACAGATAGTGTGATCTCAGCTGAATATTATCCAGTAGAA 1413
Db
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY
1414 GATGNTGTTATGGAAGCGCAGAGACTCATTTGAGTCAGTGCTGATTTCTGTGAGCGCACTA 1473
Db
421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY
1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCATGACATCTTTCATGTTTTT 1533
Db
441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY
1534 CCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGGCTTGAATGCAAAACAGGTTTC 1593
Db
461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY
1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAAAGATCCAGTGGT 1653
Db
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY
1654 GGGCTGCCTGCTCCCAAGTGATTTCAAGTGTCCTATCAAGAGGAGATAGCAATATCCAGT 1713
Db
501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY
1714 CGTGAATGGGAAGTTCTTTGGCCGCGCATGGATCTAATATCAAGTTGATGAAGTCAGAAG 1773
Db
521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY
1774 CTGGGTATATTTTGAAGGCCACCAAGACATCCCTTTAGAGCATCACCTGTAGTACGT 1833
Db
541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY
1834 TACGTAAATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db
561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY
1894 ATCAGTCAGCAGCTGTGACTTCTTTATAGTAAGTATAGTAAACCAGAGAAATCCACACTGT 1953
Db
581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY
1954 GTGTCCCTTTACAGCTATCAAGTCTCGAAGATGACCCACTTGGCAAAACAAAGGAATTT 2013
Db
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY
2014 TGGGCCACCAATTTTGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGCAAAATTTTC 2073
Db
621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY
2074 TCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGGCCTCATGATCTA 2133
Db
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL 660
QY
```

```
2134 CAGCCTGGAAGAAATATCTCTACTGTCTGTTTCATATATGTTGGTCTCTCCTCAGGTGCAGT 2193
Db
660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
QY
2194 TGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTTCGGCTTGAATACCTTAGCCTCTCTAG 2253
Db
680 lYtyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
QY
2254 GTTATGTGTTGTAGTATAGACAAACAGGGGATCCCTGTTCACCGAGGCTTTAAATTTGAAG 2313
Db
700 lYValaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnI 720
QY
2314 GGCCTTTTAAATATAAATGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAAT 2373
Db
720 yzLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerT 740
QY
2374 ATCTAGCTTCTCGATATGATTTTCACTTAGATCTGTGGGCATCCACGGCTGTGCTCT 2433
Db
740 yzGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
QY
2434 ATGGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTGAGATATCTTCAGGCTTGCTA 2493
Db
760 leAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrM 780
QY
2494 TTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGTACACGGAACGTTATA 2553
Db
780 etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG 800
QY
2554 TGGGTCACTCCCTGACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGCAG 2613
Db
800 luLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
QY
2614 AAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATG 2673
Db
820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA 840
QY
2674 TCCATTTTGACATACACCATGATATATTACTGATTTTTTAGTGGGCTGGAAAGCCATATG 2733
Db
840 splLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHist 860
QY
2734 ATTTACAGATCTATCTCTCAGGAGAGACACACATAAGAGTTCTCTGAATCGGAGAACATT 2793
Db
860 yzGluLeuHisLeuLeuHisTyrLeuGlnLysLeuGlnLysArgIleAlaAlaLeuL 880
QY
2794 ATGAACCTGCATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTAATTCGTGCTCTAA 2853
Db
880 ysvallie 882
QY
2854 AAGTGATA 2861
Db
```

RESULT 10

```
US-10-982-512-12
; Sequence 12, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ'ID'S: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-10-982-512-12

Alignment Scores:

Pred. No.: 0 Length: 4829
Score: 4680.00 Matches: 882
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 99.8% Indels: 2
DB: 9 Gaps: 0

US-10-825-632-1 (1-882) x US-10-982-512-12 (1-4829)

Qy 1 MetAlaAlaMetGluThrGluGluLeuGluValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGGAGATATTGAAACATCGGACTGT 273
Qy 21 GluGluAenIleGluSerGlnAAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATATTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACAGAAAAATATCATGGCTTACATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGACCAACATGATTTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAACTATTACCTTGCATGCTCTGGTGTGAAACAGAGAAAATACACTGTTTATTCT 513
Qy 101 GluIleProLysThrIleAenArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTCCAAACATCATATAGACAGCAGCTTAAATGCTCTCTTGGAGCCCTTTTGG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 574 GATCTTTTCAGGCACACCTGGACTATGGAATGTATCTCGAAGAAAGAACTATTAAAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAACCAATGGAAACAGTCGGAAATTTGCTTTCAGATATACCAAGGAAGTGGGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGCTAAAAAGATGGAGGCCCAAGGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTACGCCAACCAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTriPileAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT 873
Qy 221 TriPileSerAenIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAACATCTGAACCAAGAGAAGAAGAGACTCACTTATGTGCAACATGAGCTA 933
Qy 241 AlaAenMetGluGluAAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluGlu 260
Db 934 GCCAACATGGAAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAGAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTyrTrpCysProLysAlaGluThrProSerGlyGly 280
Db 994 TTTGATAGATATCTCGGCTATTTGGTGTGTCCAAAAGCTGAAACACTCCCAAGTGGTGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleIleHisVal 300
Db 1054 AAAATTTCTTAGAATTTCTATATGAAGAAAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320

Db 1114 ACATCCCTATGTTGGAAACAAGGAGGCGAGATTCATTCCGTTATCTCTAAAACAGGTACA 1173
Qy 321 AlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCBAATCTTAAGTCACCTTTTAAGATGTCAGAAATATATGATGCTGCTGAAGGAAGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCATAGATAGGAACCTAATTCACCTTTTGAGATTCTATTGAGAGGTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGACTGGATGAGCTCTGAGGAGAAATATCTTGGTCCATCTCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCCAAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATTATCCAGTAGAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAAGGACGAGACTCAITGAGTCAGTCGCTGATTCGTGACGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTriPileAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1534 CCCCAAGTCACGAAGAGGAAATTTAGTATTATTTTGGCTCTCAATGCAAAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATCAAAATTTACATCTATTTTAAAGGAAGCAAAATATAACGATCCAGTGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGCGTCTCTCCAGTGATTTCAAGTGCTCTCAAGAGAGATAGCAATTACCAT 1713
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAenIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTCTTGGCCGCGATGGATTAATATCCAGTTGATGAAGTCAGAGG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLysLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCCCTTTAGAGCATCACCTGACGTAGTCAST 1833
Qy 541 TyrValAenProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TAGCTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC 1893
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCACTGTGACTCTTTTATTAAGTATAGTAACTAACCAAGAAATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGGCCACCAATTTGGATTTCAGCAGTCTCTCTCTGACTATACTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTTCTGGATTTCATATGTATGGGATGCTCTCAAGGCTCATGATCTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL 660
Db 2134 CAGCTTGAAGAAATATCTTACTGTCTGTTTATATATGGTGGTCTCTCTCAGGTGCAST 2193
Qy 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuG 680

QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
DB	1234	ATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGAGATTCTATTATGGAAGGAGTTGAA	1293
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrPheIleLeuLeuAsp	380
DB	1294	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAAATATGCTTGGTCCATCCTACTAGAT	1353
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
DB	1354	CGCTCCAGACTCCCTCAGATAGTGTGATCTCACCTGATATTATTATCCAGTAGAA	1413
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
DB	1414	GATGATGTTATGGAAAGGCGAGACTCATTTGAGTCACTGCTGATCTCTGTGAGCCACTA	1473
QY	421	IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe	440
DB	1474	ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTCATGTGTTT	1533
QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
DB	1534	CCCCAAAGTCACGAAGAGAAATTGAGTTTATTTTTCCTCTGAAATGCAAAACAGGTTC	1593
QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLeuGluSerLysTyrLysArgSerSerGly	480
DB	1594	CGTCATTTTATACAAAATATACATCTATTTTAAAGAAAGCAAAATATAAAACGATCCAGTGGT	1653
QY	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
DB	1654	GGGCTCGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAAAGAGAGATAGCAATATACCAAGT	1713
QY	501	GlyGluTyrGlnValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
DB	1714	GGTGAAATTTTGAAGGACCAAAAGACTTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
QY	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuValValSer	540
DB	1774	CTGGTATATTTTGAAGGACCAAAAGACTTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
DB	1834	TACGTAAATCCCTGGAGAGGTGACCAAGGCTGACTGCGTGGCTACTCACATTTCTTGCTGC	1893
QY	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
DB	1894	ATCAGTCAGCATGTGACTTCTTTATAGTAAGTATAGTAACCAAGAGAAATCCACACTGT	1953
QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
DB	1954	GTGTCCCTTTACAAGCTATCAAGTCTCAAGATGACCCCAACTTGCAAAACAAAGGAATTT	2013
QY	601	TyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
DB	2014	TGGGCCACCATTTTGGATTCACT-----	2036
QY	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
DB	2036	-----	2036
QY	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
DB	2037	-----CCTCAGGTGCAAGTTG	2051
QY	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
DB	2052	GTGAATAATCGGTTTAAAGAGTCAAGTATTTTCCCTTGATACCTTAGCTCTCTAGGT	2111
QY	681	TyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly	700
DB	2112	TATGTGGTTGTAGTAGACAAACAGGGGATCCTGTCAACGAGGGCTTAATTTGAAGGC	2171
QY	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
DB	2172	GCCTTTAAATATATAAATGGTCAATAGAAATGACGATCAGGTGAAGGACTCCAATAT	2231
QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyLysHisGlyTyrSerTyr	740
DB	2232	CTAGCTTCTCGATATGATTTTCATTGATTTAGATCTGTGGGCATCCACGGCTGGTCTCTAT	2291
QY	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
DB	2292	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTCCAGATATCTTCAGGGTTCCTATT	2351
QY	761	AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMet	780
DB	2352	GCTGGGGCCCGACGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2411
QY	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
DB	2412	GGTCACCTGCACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2471
QY	801	LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
DB	2472	AAGTTCCTCTGAAACCAATCGTTTACTTCTTACATGTTTCTCGATGAGAATGTC	2531
QY	821	HisPheAlaHisThrSerIleLeuLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
DB	2532	CATTTTGACATACCATCATATTACTAGTCTTTTGTAGTGAGGCTGGAAGGACATATGAT	2591
QY	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
DB	2592	TTACAGATCTATCTCTCAGGAGACACACAGCATTAAGAGTTCTCTGAATCGGAGACATTAT	2651
QY	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
DB	2652	GAACTGCACTCTTTTGCACTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTAATA	2711
QY	881	ValIle 882	
DB	2712	GTGATA 2717	
RESULT 12			
US-10-982-512-22			
; Sequence 22, Application US/10982512			
; Publication No. US20050059081A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/10/982,512			
; CURRENT FILING DATE: 2004-11-05			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 22			
; LENGTH: 4685			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-982-512-22			
Alignment Scores:			
Pred. No.:	0	Length:	4685
Score:	4385.50	Matches:	834
Percent Similarity:	94.6%	Conservative:	0
Best Local Similarity:	94.6%	Mismatches:	1
Query Match:	93.3%	Indels:	48
DB:	9	Gaps:	1

US-10-825-632-1 (1-882) x US-10-982-512-22 (1-4685)

QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 214 ATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTTGAAACCTGGGACTGT 273
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 274 GAGGAGAAATATTGAATACAGATCGGCTTAAATTTGAGAGCTTTTATGTTGAGCGGTAT 333
QY 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCCTGGAGTCAGCTTAAAAGCTGCTTGGCCGATACCAAGAAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAAATCTATTACCTTGGCCATGCTGGTGAGAACAGAGAAATACACTGTTTATTCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTCCTCAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTGGAGCCCTCTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 574 GATCTTTTCAGCACAACCTGGACTATGGAATGATTTCTCGAGAGAAGAACTATTAAAG 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAAACCGATTTGGAACAGCTCGGAATTCCTTTACGATTATCACCAAGAGAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATCACGTAAAGATGAGGGCCACAAAGA 753
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTAGCAACAACCTTTTAAGGCCCAATCTAGTGAACACTAGTTGTCCCAACATACGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 814 GATCCAAAATATTATGCCCTGCTGATCCAGACTGGATTCCTTTATACATAGCAACGATAT 873
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
DB 874 TGGATATCTAAATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCAATGAGCTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 934 GCCAATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATCTGGCTATTGGTGTGTCGAAAGCTGAACAACTCCACAGTGTGT 1053
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
DB 1054 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTATCATGTT 1113
QY 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTTATGTTGGAACCAAGGAGGCGAGATTCATTTCCGTTATCTCTAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 1174 GCNAATCTTAAGTCACTTTTAAGNTGTCAAGAAATATGATGCTGAGGAAGAGATC 1233
QY 341 IleAspValIleAspLysLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTTGAAGAGTTGAA 1293

QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTGCCAGAGCTGGATGGACTCTGAGGAAATATGTTGGTCCATCTACTAGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
DB 1414 GATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1474 ATTATCTATGAGAAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1534 CCCAAAGTCACGAAGAGGAAATTTGAGTTATTTTGGCTCTGAAATGCNAACAGAGTTTC 1593
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
DB 1594 CGTCATTTATACAAAATTACATCTATTTAAAGGAAGCAATATATAACGATCCAGTGT 1653
QY 481 GlyLeuProAlaProSerAspPheLysCysProLysLysGluGluIleAlaIleThrSer 500
DB 1654 GGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTAACCAT 1713
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
DB 1714 GGTGAATGGGAAGTTCTTGGCCGCGATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAGATAGTATACCAAGAGAAATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCCCTTTACAGCTATCAAGTCTGAGATGACCCCACTTGCAAAACAAAGAGAAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
DB 2014 TGGGCCACCATTTTGGATTTCAGT----- 2036
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2036 ----- 2036
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 2037 -----CCTCAGGTGCGATTG 2051
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB 2052 GTGAATATCGGTTTAAAGAGTCAAGTATTTCCGTTGATATACCTAGACCTCTCTAGGT 2111
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
DB 2112 TATGTGGTTGTAGTAGACAAACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGGC 2171
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGluTyr 720
DB 2172 GCCTTTAAATATATAAATGGGTCAATAGAAATTTGCGATCAGGTGGAGGAGTCAATAT 2231

QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
Db	2232	CTAGCTTCTCGATATGATTTTCATGTAGCTGFGTGGGCATCCACGGCTGGTCCCTAT	2291
QY	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2292	GGAGGATACCTCTCCCTGATGGCATTAATGCGAGGTCAGATATCTTCAGGGTTGCTATT	2351
QY	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2352	GCTGGGGCCCGACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2411
QY	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	2412	GGTCACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2471
QY	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2472	AAGTTCCCTCTGNAACCAATCGTTTACTGCTTACATGGTTTCTCGATGAGAAATGTC	2531
QY	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2532	CATTTTGACATACAGTATATTAATGAGTTTCTAGTGGGCTGGAAGCCATATGAT	2591
QY	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2592	TTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTTGAAATCGGAGAACATTAT	2651
QY	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
Db	2652	GAACATGCAATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTGCTGCTCAAAA	2711
QY	881	ValIle 882	
Db	2712	GTGATA 2717	
RESULT 13			
; Sequence 20, Application US/09976674			
; Patent No. US20020115843A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976.674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 20			
; LENGTH: 4676			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-976-674-20			
Alignment Scores:			
Pred. No.:	0	Length:	4676
Score:	4385.00	Matches:	831
Percent Similarity:	94.2%	Conservative:	0
Best Local Similarity:	94.2%	Mismatches:	1
Query Match:	93.3%	Indels:	51
DB:	3	Gaps:	1
US-10-825-632-1 (1-882) x US-09-976-674-20 (1-4676)			
QY	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	214	ATGGCAGCAGCAATGGAAACAGAACACAGCTGGGTGTGAGATATTGGAATCTGGCGGCTGT	273


```
Db 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCCTGAATATATTTATCCAGTAGAA 1413
Qy 401 AspAspValMetGluAtrGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAGGACAGAGACTCATTTAGTCAGTCGCTGATTCCTGAGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysIysThrGlyPhe 460
Db 1534 CCCCAAGTFCAGAGAGAGAAATTTAGTTATTTTGGCTCTGAATGCAAAAACAGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 1654 GGCTGCTGCTCCCAAGTATTTCAAGTGTCTTATCAAGAGAGATAGCAATTTACCACT 1713
Qy 501 GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAATATCCAACTTGATGAAGTCAGAAG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAGT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TACGTAAATCTGGAGAGGTGACCAAGCTGACTGACCGTGGCTACTCACATCTTGTCTG 1893
Qy 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCAGCTGTGACTTCTTTAAGTAAGTATAGTAAACCAAGAAATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProLysAspProThrCysIysThrLysGluPhe 600
Db 1954 GTGTCCCTTTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAAACAAAGGAATT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGCCACCATTTTGATTCAGCAGCTCTCTCTGACTATCTCTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA 2133
Qy 641 GlnProGlyLysGlyTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCCTGGAAGAAATATCTCTACTGTCTGTGTATATATGTTGGTGGTCG----- 2180
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2180 ----- 2180
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2180 ----- 2180
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2181 -----GGTCAATAGAAATTCAGCATCAGCTGGAAGGACTCCCAATAT 2222
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2223 CTAGCTTCTCGATATGATTTTCACTTACCTGTGGGCATGCCAGCGCTGGTCTTAT 2282
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
```

```
Db 2283 GGAGGNTACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATT 2342
Qy 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2343 GCTGGGGCCCCAGTCACTCTCTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2402
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2403 GGTCAACCTTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2462
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2463 AGTTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAATGTC 2522
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2523 CATTTTGCACATACCAGTATATTTACTGAGTTTTTTAGTGGGGCTGGAAGCCATATGAT 2582
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2583 TTACAGATCTATCTCCAGAGAGACACAGCATAAGAGTTCTGAATCGGAGAACATTAT 2642
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2643 GAATCGATCTTTTGCACTACTCTTCAGAAAACCTTGGATCAGCTATTGCTGCTCTAAAA 2702
Qy 881 ValIle 882
Db 2703 GTGATA 2708
```

RESULT 14

```
US-10-982-512-20
; Sequence 20, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-20
```

```
Alignment Scores:
Pred. No.: 0 Length: 4676
Score: 4385.00 Matches: 831
Percent Similarity: 94.2% Conservative: 0
Best Local Similarity: 94.2% Mismatches: 1
Query Match: 93.3% Indels: 51
DB: 9 Gaps: 1
```

US-10-825-632-1 (1-882) x US-10-982-512-20 (1-4676)

```
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAATATGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATTTGTTGAGCGGTAT 333
```



```

Db 2343 GCTGGGGCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGNACGTTATATG 2402
Qy 781 GlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2403 GGTCACTCCAGCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2462
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAnVal 820
Db 2463 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGGAATGTC 2522
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2523 CATTTGGACATACAGATATATTACTGATTTTTTTAGTGGGCTGGAAGCCATATGAT 2582
Qy 841 LeuGlnIleTyrProGlnGlnuArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2583 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTGAAATCGGAGAACATTAT 2642
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGlnuAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2643 GAACGTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTATTGCTGCTCAAAA 2702
Qy 881 ValIle 882
Db 2703 GTGATA 2708

RESULT 15
US-10-956-157-2177
; Sequence 2177, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2177
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-2177

Alignment Scores:
Pred. No.: 0 Length: 2830
Score: 4118.00 Matches: 782
Percent Similarity: 88.7% Conservative: 0
Best Local Similarity: 88.7% Mismatches: 0
Query Match: 87.6% Indels: 100
DB: 9 Gaps: 1

US-10-825-632-1 (1-882) x US-10-956-157-2177 (1-2830)
Qy 1 MetAlaAlaMetGluThrGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAACACAGACGCTGGGTGTGAGATATTTGAACTCGCGACTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACCAAGAAATATCATCATGCTCATGATG 393
Qy 61 AlaLysAlaProHisAsnPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATTCAGATGACCTCATTTCA 453
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAnThrLeuPheTyrSer 100

```

```

Db 454 GACAGAAATCTATTACTTGGCCATGCTCTGTGTGAGACACAGAGAAATAACACTGTTTATCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGGCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCAGAGAAGAACTATTATGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAAACGATTTGGAACAGTTCGGAATTCCTTACGATTATCACCAGGAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCGGTAGTGGATTTATCAGGTAAAGATGAGGCGCCACAAGGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTTACGCAACACCTTTTAAAGGCCAATCTAGTGGAACTAGTTGTCCTCCACATACGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAlaPheIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAAATATATGCCCTGCTGATCCAGACTGGATTGCTTTTATACATACGACGATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAACATCGTAACACAGAGAGAGGAGACTCACTTATGTGCACAATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTTTGTCTCCAAGAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTTTGGCTTATTTGGTGGTGTCCAAAGCTGAACAACTCCACGFGTGGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 1054 ABAATTTCTAGAAATCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCTATGTTGGAAACAAGGAGGCGAGATTCAATCCGTTTATCTCTAAACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAAAATCCTTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGATGCTGAAGGAGGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCATAGATAGGAAGAACTAATTCACCTTTTGAGATTCATTGGAAGGATGGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGAGCTGGATGGACTCCTCGAGGAAAATATGCTGCTCCATCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuLeuSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCAGACTGCGCTACAGATAGTGTGATCTCACCTGAAATATTATTATCCCAAGTGA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAGGAGGAGACTCAATGATGCTGCTGCTGATCTGTGACGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAGAAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Qy 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460

```

```

Db 1534 CCCAAGTACGAGAGGAATTTGAGTTATTTTGGCTCGAATGCCAAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATACAAATTTACATCTATTTAAAGGAAGCAATAATAACGATCCAGTGGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGCTGCCTGCTCCAGTGATTTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATACCAGT 1713
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
Db 1714 GGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAAGTTGATCAATCAGAGG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTAGTACAGT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TACGTAAATCTCGAGAGGTGACNAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCACTGTGACTCTCTTATAGTATAGTATACCAAGAGAAATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysIysThrLysGluPhe 600
Db 1954 GTGTCCTTTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTGACTATACTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTACATTTGATGGATGCTCTACAAAGCCTCATGATCTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCCTGGAAGAAATATCTCTACTGTCTGTCTATATATGTTGCTCCTCAG----- 2184
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2184 ----- 2184
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2184 ----- 2184
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2184 ----- 2184
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2184 ----- 2184
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2185 -----GTTCGTATT 2193
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2194 GCTGGGGCCCCAGTCACCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2253
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2254 GGTCAACCTTGACCAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2313
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2314 AGTTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGATGTC 2373

```

```

Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2374 CATTTGCACATACCAGTATATTACTGAGTTTTTTAGTCAGGGCTGGAAAGCCATATGAT 2433
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleAsnValProGluSerGlyGluHisTyr 860
Db 2434 TTACAGATCTATCTTCAGGAGACACACAGCATTAAGAGTTCTCTGAATCCGGAGAACATTAT 2493
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2494 GAACCTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTTTCTCTTAAAA 2553
Qy 881 ValIle 882
Db 2554 GTGATA 2559

```

Search completed: April 15, 2006, 01:36:23
Job time : 1586 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

- OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 01:11:56 ; Search time 855 Seconds
(without alignments)
4155.026 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAAAMTEQLGVIEFETADC.....HLHYLQENLGSRIIAALKVI 882

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9281099 seqs, 2013915447 residues

Total number of hits satisfying chosen parameters: 18562198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=rlp
-Q=/abs/ABSSWEB.spool/US10825632/runat_14042006_092613_5040/app_query.fasta_1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs805P
-USER=US10825632 @CGN 1.1 991 @runat_14042006_092613_5040 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /SID55/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /SID55/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /SID55/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /SID55/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
8: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq2:*
9: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq3:*
10: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq1:*
11: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
12: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
13: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
14: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq5:*
15: /SID55/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	4700	100.0	2649 14	US-11-151-601-21
2	4700	100.0	3143 14	US-11-151-601-19
				Sequence 21, Appl
				Sequence 19, Appl

3	651	13.9	2238	11	US-11-079-463-1186	Sequence 1186, Ap
4	551	11.7	1884	11	US-11-079-463-2536	Sequence 2536, Ap
5	529	11.3	2217	11	US-11-208-288-3	Sequence 3, Appl
6	529	11.3	3332	11	US-11-208-288-1	Sequence 1, Appl
7	529	11.3	3407	9	US-10-501-035-34	Sequence 34, Appl
8	525	11.2	2301	8	US-10-522-789-1	Sequence 1, Appl
9	517.5	11.0	2283	11	US-11-208-288-5	Sequence 5, Appl
10	490	10.4	2778	9	US-10-932-182A-5649	Sequence 5649, Ap
11	490	10.4	2778	9	US-10-932-182A-5649	Sequence 5649, Ap
12	478.5	10.2	4852	14	US-11-136-527-2130	Sequence 2130, Ap
13	454	9.7	2814	14	US-11-245-147-168	Sequence 168, Ap
14	454	9.7	2814	14	US-11-186-284-54	Sequence 54, Appl
15	436.5	9.3	2457	9	US-10-932-182A-1107	Sequence 1107, Ap
16	436.5	9.3	2457	9	US-10-932-182A-1107	Sequence 1107, Ap
17	433	9.2	2819	14	US-11-136-527-2716	Sequence 2716, Ap
18	278.5	5.9	535	11	US-11-226-869-428	Sequence 428, Ap
19	252.5	5.4	1875	11	US-11-079-463-2282	Sequence 2282, Ap
20	215	4.6	1346	6	US-09-925-065A-669313	Sequence 669313, A
21	211.5	4.5	600	14	US-11-136-527-6812	Sequence 6812, Ap
22	206.5	4.4	2214	11	US-11-079-463-4059	Sequence 4059, Ap
23	200	4.3	609	6	US-09-925-065A-743558	Sequence 743558, A
24	200	4.3	609	6	US-09-925-065A-743558	Sequence 743558, A
25	194	4.1	378	9	US-10-932-182A-81332	Sequence 81332, A
26	194	4.1	378	9	US-10-932-182A-81332	Sequence 81332, A
27	181.5	3.9	354	9	US-10-932-182A-173537	Sequence 173537, A
28	181.5	3.9	354	9	US-10-932-182A-173537	Sequence 173537, A
29	156	3.3	2384	14	US-11-136-527-2023	Sequence 2023, Ap
30	130.5	2.8	2112	9	US-10-932-182A-81333	Sequence 81333, A
31	130.5	2.8	2112	9	US-10-932-182A-81333	Sequence 81333, A
32	127	2.7	612	6	US-09-925-065A-818446	Sequence 818446, A
33	126.5	2.7	4344	9	US-10-932-182A-3542	Sequence 3542, Ap
34	126.5	2.7	4344	9	US-10-932-182A-3542	Sequence 3542, Ap
35	125.5	2.7	5493	9	US-10-932-182A-81749	Sequence 81749, A
36	125.5	2.7	5493	9	US-10-932-182A-81749	Sequence 81749, A
37	125	2.7	4257	9	US-10-932-182A-4976	Sequence 4976, Ap
38	125	2.7	4257	9	US-10-932-182A-4976	Sequence 4976, Ap
39	124	2.6	9023	8	US-10-995-561-134	Sequence 134, App
40	123.5	2.6	1694969	7	US-10-506-454-1690	Sequence 1690, Ap
41	121.5	2.6	3464	8	US-10-995-561-465	Sequence 465, App
42	119.5	2.5	3387	14	US-11-091-883-183	Sequence 183, App
43	119.5	2.5	3439	14	US-11-000-688-851	Sequence 851, App
44	118.5	2.5	3468	14	US-11-098-686-9205	Sequence 9205, Ap
45	118.5	2.5	3729	8	US-10-453-372-1135	Sequence 1135, Ap

ALIGNMENTS

RESULT 1

US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US2006003413A1

GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark

; APPLICANT: Welch, Nadine

; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF

; FILE REFERENCE: MPI00-054PIRCPIOMNIDVIM

; CURRENT APPLICATION NUMBER: US/11/151,601

; PRIOR FILING DATE: 2005-06-13

; PRIOR APPLICATION NUMBER: US 10/170,789

; PRIOR FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: US 09/797,039

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: PCT/US01/06525

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 60/186,061

; PRIOR FILING DATE: 2000-02-29

; PRIOR APPLICATION NUMBER: US 09/882,166

```

; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ IDS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-151-601-21

Alignment Scores:
Pred. No.: 0 Length: 2649
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0

US-10-825-632-1 (1-882) x US-11-151-601-21 (1-2649)
QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 1 ATGGCAGCAGCAATGGAACAGAACAGCAGCTGGGTGTTGAGATATTGAAACTCGCGACTGT 60
QY 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 61 GAGGAGATATTGAAATCATCAGCATCGGCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 120
QY 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 121 TCTCGAGTCAGCTTAAAAAGCTGCTTGGCGATACACAGAAATATCATGGCTACATGATG 180
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 181 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 240
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAenThrLeuPheTyrSer 100
Db 241 GACGAAATCTATTATCTTCCCATGCTGTGTGAGAACAGAGAAATACACTGTTTATTCT 300
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 301 GAAATTTCCAAACATCATATAGACAGCAGCTTATATGCTCTTTGGAAGCCTCTTTTG 360
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 361 GATCTTTTTCAGGCAACATCGGACTATGGAATGATATCTCGAAGAAAGAACTATTAAAG 420
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 421 GAAAGAAAACGGATCGGAACAGTCGGAATTCCTTACGATATATACCAAGAGGTGGA 480
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 481 ACATTTCTGTTCACGCGGTAGTGGAAATTTATACGTPAAAGATGAGGCGGCACAGGA 540
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 541 TTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCGCCAACATACGATG 600
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 601 GATCCAAAAATTATGCTCTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATATT 660

```

```

QY 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
Db 661 TGGATATCTAACATCGTAAACCAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 720
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 721 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAAAGAA 780
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 781 TTTGATAGATATTCTGGCTATTGCTGTGTCCAAAGCTGAAACCAACTCCCAAGTGTGT 840
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 841 AAAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTCTGTT 900
QY 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 901 ACATCCCCCTATGTTGGAACAAGAGGCGCAGATTCATTCCTGTTATCTCTAAACAGGTACA 960
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 961 GCMAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGGAAGAGGATC 1020
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1021 ATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATCTCTATTTGAAGGAGTTGAA 1080
QY 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1081 TATATTGCGACAGCTGGATGGACTCTCTGAGGAAAAATATGCTTGGTCCATCTCTACTAGAT 1140
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1141 CGCTCCAGACTCGCTCGCAGATAGTGTTCATCTCACCTGAAATTAATTTATCCAGTAGAA 1200
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1201 GATGATGTTATGGAAGAACAGAGACTCATTTGAGTCAGTCGCTGATCTGTGAGCGCACATA 1260
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1261 ATTATCTATGAAGAAACACAGACATCTGGNTAATATCCATGACATCTTTTCATGTTTTT 1320
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1321 CCCCAAGATCAGAAAGAGAAATGAGTTTATTTTTCCTGATGATGCAAAACAGGTTTC 1380
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1381 CGTCATTTTATACAAATTTACATCTATTATTAAGGAAAGCAATATAACGATCCAGTGGT 1440
QY 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
Db 1441 GGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTCAAGAGGAGATAGCAATATCCAGT 1500
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1501 GGTGAATGGAAAGTTCTTGGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1560
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1561 CTGGTATATTTTGAAGGCACCAAGACTCCCCCTTTAGAGCATCACCTGTAGTAGTCAGT 1620
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1621 TACGTAAATCTGTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACATTTCTTGTCTG 1680
QY 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1681 ATCAGTCAGCAGCTGTGACTTCTTTTATAGTAGTATAGTAACCAAGAAATCCACACTGT 1740

```

```
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspPheProThrCysLysThrLysGluPhe 600
Db 1741 GTGTCCTCTTTACAGCTATCAAGTCTCAAGATCACCACTTGCAAAACAAAGGAATTT 1800
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1801 TGGGCCACCATTTTGGATTACAGGCTCTCTTCTGACTATCTCTCCAGAAATTTTC 1860
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1861 TCTTTTGAAGTACTACTGGATTATCATTTGATGGATGCTCTACAAGCCTCATGATCTA 1920
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1921 CAGCCTCGAAGAAATATCTACTGCTGTGTTTCATATATGGTGGCTCTCAGGTGCAAGTTG 1980
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1981 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGT 2040
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2041 TATGTGTTGTAGTAGACACAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGGC 2100
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2101 GCCTTTAAATATAAATGGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATAT 2160
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2161 CTAGCTTCTCGATATGATTTTCATTCATGCTGTGGGCATCCACGGCTGGTCTCTAT 2220
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2221 GGGAGATACCTCTCCCTGATGTCATTAATGCAGAGGTCAATATCTTCAGGGTTGCTATT 2280
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2340
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2341 GGTCAACCTCGACAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2400
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2401 AAGTCCCTCTGACACCAATGCTTTACTGCTCTTACATGTTTCTCGGATGAGATGTC 2460
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2461 CATTTTGCACATACCACTATATTAAGTATTTTCTAGTGGGCTGGAAAGCCATATGAT 2520
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2521 TTACAGATCTATCTCCTCAGGAGACACAGCATTAAGATTCCTGAATCGGAGAACATTAT 2580
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2581 GAACTGCATCTTTTGCACTACCTTCAAGAAACCTTGGATTCAGGTATTTGCTGCTCTAAA 2640
Qy 881 ValIle 882
Db 2641 GTGATA 2646
```

RESULT 2

```
US-11-151-601-19
; Sequence 19, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
```

```
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229) ... (2874)
US-11-151-601-19
```

Alignment Scores:

Pred. No.:	0	Length:	3143
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	14	Gaps:	0

US-10-825-632-1 (1-882) x US-11-151-601-19 (1-3143)

Qy	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	229	ATGGCAGCAGCAATGGAAACACAGACGCTGGGTGTGAGATATTGAAACTGCGGACTGT	288
Qy	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Db	289	GAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT	348
Qy	41	SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Db	349	TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACACAGAAATATCATGGCTACATGATG	408
Qy	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Db	409	GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTC	468
Qy	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Db	469	GACAGAAATCTATTACCTTCCCATGCTGTGTGAGAACAGAGAAATACACTGTTTTATTCT	528
Qy	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTyrLysProLeuLeu	120
Db	529	GAATTTCCCAAACTATCAATAGACAGCAGTCTTTAATGCTCTCTTGGAGCCCTCTTTTG	588

Qy	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140	Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluLeuAlaIleThrSer	500
Db	589	GATCTTTTTCAGGCAACACTGGACTATGGAAATGATATTCGAGAAAGAACTATTAAAGA	648	Db	1669	GGGCTGGCTGCTCCAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTACCAGT	1728
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	Qy	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	649	GAAGAAGAAACGCAATGGACACAGTCGGAAATGCTTCTTACGATTAATCAACCAAGGAAGTGA	708	Db	1729	GGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCAAGTTGATAGTCAGAGG	1788
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
Db	709	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTAATACGTAATAAGATGGAGGGCCACAAGGA	768	Db	1789	CTGGTATATTTTGAAGGCCCAAGACTCCCTTTTAGAGCATCCTCTACGTAGTCACT	1848
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	769	TTTACGCCAACCACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACCGATG	828	Db	1849	TACGTAATACTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTCTCTC	1908
Qy	201	AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle	220	Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	829	GATCCAAATTAATGCCCTGCTGATCCAGACTGGATTCCTTTATACATAGCAACGATATT	888	Db	1909	ATCAGTCAGCACTGTGACTCTTTTATTAAGTAAGTATAGTAACCAAGAAATCCACTGT	1968
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu	240	Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	889	TGGATATCTTAACATCGTAACCGAGAGAAAGAGAGACTCACTTATGTGCACATGAGCTA	948	Db	1969	GTGTCCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTT	2028
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260	Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
Db	949	GCCAACTGGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCAGAGAA	1008	Db	2029	TGGGCCACCATTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2088
Qy	261	PheAspArgTyrSerGlyTyrTrpTyrProLysAlaGluThrThrProSerGlyGly	280	Qy	621	SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	1009	TTTGTATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGT	1068	Db	2089	TCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGCTCATGATCTA	2148
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Qy	641	GlnProGlyLysGlyTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	1069	AAAAATCTTAGAAATTTCTATATGAGAAATGATGAATCTGAGGTGAAATTAATCATGTT	1128	Db	2149	CAGCTCGGAAGAAATATCTTACTGTGCTGTTCATATATGCTGCTCTCAGGTGCAGTTG	2208
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	1129	ACATCCCTATGTTGAAACCAAGAGGGCCAGATTCAATCCGTTATCTCTAAAAACAGGTACA	1188	Db	2209	GTGAATAATCGGTGTAAAGGAGTCAAGTATTTCCGTTGATATCCCTTAGCTCTCTAGGT	2268
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Qy	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	1189	GCAAACTCTTAAAGTCACATTTTAAAGATGTCAGMAATATGATGATGCTGAAGGAAGATC	1248	Db	2269	TATGTGTTGTAGTGATAGAACACAGGGGATCCTGTCAACAGGGCTTAATTTTGAAGGC	2328
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	1249	ATAGATGTCATAGATAAGGAACTAATTCACCTTTTCAGATTTCTATTGGAAGGAGTTGAA	1308	Db	2329	GCCTTTTAAATATAAAATCGGTCAATAGAAATTCAGCATCAGGTGGGAAGGACTCCAATAT	2388
Qy	361	TyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Db	1309	TATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1368	Db	2389	CTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGGGCATCCACGGCTGGTCTAT	2448
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	1369	CGCTCCAGACTCGCTCGAGATAGTGTGATCTCACTGATATTAATTTATCCAGTAGAA	1428	Db	2449	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTGGCTATT	2508
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Qy	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	1429	GATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGGCTGATCTGAGCCCACTA	1488	Db	2509	GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2568
Qy	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	1489	ATTATCTATGAAGAACACAGACATCTGGATAAATATCCATGACATCTTTCAATGTTT	1548	Db	2569	GGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGACGAA	2628
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	Qy	801	LysPheProSerGlnProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	1549	CCCCAAGTCAGGAAGAGAAATTGAGTTTATTTTTCCTCTGTAATGCCAAACAGGTTTC	1608	Db	2629	AAGTTCCCTCTGACCAANAATCGTTTACTGTCTTACATGGTCTTCTGGATGAGAATGTC	2688
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	Qy	821	HisPheAlaHisThrSerIleLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	1609	CGTCAATTAACAAAATACATCTATTTTAAAGGAAAGCAATATAAACATCCAGTGGT	1668	Db	2689	CATTTGCAATACCAAGATATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGAT	2748
Qy				Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860

Db	1000	TACCTGATGTGAACCAATACAAC	TCCAGACTGTTCCGAGAAGTATCAA	TATCATCGATGTG	1059			
Qy	587	SerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeuAsp	606					
Db	1060	CANAAGTGTAAGC-----	GTCAAATCTGCTACA	1089				
Qy	607	SerAlaGlyProLeuProAspTyrThrProPro-	-GluilePheSerPheGluSer	624				
Db	1090	GCAGCGGATCCTTTCCCGGCTACAAGATGCCGGGTATCGAGACAGGTACCATCAAAGCA	1149					
Qy	625	ThrThrGlyPheThr---	LeuTyrGlyMetLeuTyrLysPheProHisAspLeuGlnProGly	643				
Db	1150	GCAGATGCGAAGCGGACTTATTACCGATTATAATAAAACCGCGCGATTTCGATCCGAAC	1209					
Qy	644	LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn	663					
Db	1210	AAGAAATACCCGGCTATCGTATACGTATACGGCGGTCTCACGCACAATGGTCACCAAC	1269					
Qy	664	ArgPheLys--GlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal	682					
Db	1270	GGCTGGCAGAACGGTGCAGAGCTGGGATATC---TATATGCCCAACAAGGTTTACT	1326					
Qy	683	ValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe	702					
Db	1327	ATGTTTACCGTAGACGAGCGTGAAGCAGCAATCGGGAGCTCGATTTTGAGAATGTTACT	1386					
Qy	703	LysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrIleuAla	722					
Db	1387	TTCCGCCAGTTGGGAATCGAAGAGGAAGACAGCAGGTGAAGGAACCGCAATTCCTGAAA	1446					
Qy	723	SerArgTyrAppPheIleAspLeuaspArgValGlyIleHisGlyTrpSerTyrGlyGly	742					
Db	1447	AGC---CTCCCCACGTGCGTGAAGACCGTATCGAGATACGGCTGGAGCTTTGGCGGT	1503					
Qy	743	TyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGly	762					
Db	1504	CACATGACCACTGCCCTTCTCTCTCGGTATCCCGAGATATTTAAGTGGCGGTGGCGGC	1563					
Qy	763	AlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHis	782					
Db	1564	GGTCCCGTCATTGACTGGGGGTATTACGAAGTGATGTACGGAGAACGCTATATGGATACC	1623					
Qy	783	ProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhe	802					
Db	1624	CCGAAAGCAATCCGAAGAGTATAAAGAGGTGAACCTGAAAAAACCTGCAGCGCAACCTG	1683					
Qy	803	ProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluasnValHisPhe	822					
Db	1684	-----AAGAGCATCTCATCATCATCCAGATGACACGACCACTGGCGTGGCG	1734					
Qy	823	AlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGln	842					
Db	1735	CAGCACACCTCTCGTTTATGAAGCGTGCATCGATGCCCGACGCTATCCGACCTCTTC	1794					
Qy	843	IleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGlu-----His	859					
Db	1795	ATCTACCCGTGCCACAGCACAAATGTG-----TCGGGCGCGACCGTGTACAT	1842					
Qy	860	TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu	872					
Db	1843	CTCATGAAAGATTAACCGGTTATTTCGAAGACTATTTA	1881					

RESULT 5

US-11-208-288-3

03-11-200-280-3 : Sequence 3, Application US/11208288

; sequence 3, application 00/1120
; Publication No. US20060051366A1

: GENERAL INFORMATION:

APPLICANT: CHANG, Chiwen

APPLICANT: CHANG, CHITWELL
: TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF

: TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION

FILE OF INVENTION: ANGLO
FILE REFERENCE: 39533-0001

FILE REFERENCE: 39533-0001
CURRENT APPLICATION NUMBER: IIS/11/208.288

QY 313 PheArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIle 332
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
760 GTTCATATCCAAAGCAGAGCTGTGAATCCCACTGTAAAGTTCTTTGT-----GTA 813

QY 333 MetIleAspAlaGluGlyArgIlelleAaspValIleAaspLysGluLeuIleGlnProPhe 352
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
814 AATAACAGACTCTCTCAGCTCAGTACCACCAATGCACCTCCATACAAATCACTCGCTCTGCT 873

QY 353 GluIleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLys 372
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
874 TCTATGTTGATA---GGGATCACTACTTGCTGTGATGCATGGCA-----918

QY 373 TyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSer 392
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
919 -----ACACAAGAAGAATTCTTTCAGTGCCTC---948

QY 393 ProGluLeuPheIleProValGluAaspValMetGluArgGlnArgLeuIleGluSer 412
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
949 -----AGGAGGATTCAGAAC 963

QY 413 ValProaspSerValThrProLeuIlelleTyrrGluGluThrThrAspileTrpIleAsn 432
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
964 .TAT-----TCGGTCATGGATATTGTGACTATGATGAATCCAGTGGGAAGATGG---AAC 1014

QY 433 IleHisAspIlePheHisValPheProGlnSerHisGluGluIleGluPheIlePhe 452
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1015 TGC-----TTAGTGGCAGGCCAACACATTGAATCAGTACTACTGGCTGG 1059

QY 453 AlaSerGluCyLSyThrGlyPheArgHis-----LeutyryLys 465
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1060 GTTGGNAGATTTAGSCCTTCAGAACCTCATTTTTACCTTTGATGGTAATAGTCTTACAAG 1119

QY 466 lleThrSerIleLeuLysGluSerLysTyrrLysArgSerSerGlyGlyLeuProAlaPro 485
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1120 ATCATCAGC-----AATGAAGAAGGTTACAGACATT-----1152

QY 486 SerAppPheLysCyeProilleLysGluIleAlaIleThr-SerGlyGluTrpGluVal 505
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1153 TGCTATTTCCAANTAGATAAAAAGACTGCACATTTATTACAAAAGCACCTCGGAAGTC 1212

QY 506 LeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr-PheGlu 525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1213 ATCGGG-----ATAGAAGCTCTAACCCAGTGATTATCTATCTACATTAGT 1257

QY 526 GlyThr---LysAaspSerProLeuGluHiHisLeutyrrValIleSerTyrrValAsnPro 544
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1258 AATGAATATAAAGGAATCCAGGAGGAAGNAATCTTTATAAAATCCAATTAATTAGCTAT 1317

QY 545 GlyGluValThrArgLeuThrAspArgGlyTyr-SerHisSerCysIleSer---Gln 563
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1318 ACAAAAGTGACATGCCTC-----AGTTGTAGCTGAATCCCGGAA 1356

QY 564 HisCysAaspPhePheIleSerLysTyrr-SerAsnGlnLysAsnProHis-----Cys 580
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1357 AGTGTCAGTACTATTCTGTCTCATTCAGTAAGAGCGCAAGTATTATTCAGCTGAGATGT 1416

QY 581 -----ValserLeutyrrLysLeuSerSerProGluAaspAspProThrCysLys 596
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1417 TCCTGGTCTGGTCTGCCCTCTATACTCTACAGCAGCGCTGTAATGATAAAGGGCTGAGA 1476

QY 597 ThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyr-ThrPro 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1477 GTCCTGGAAAGAC---AANTTCAGCTTTGGATAAAATG-----CTGCAGATGTCAGATG 1527

QY 617 ProGlu-----IlePheSerPheGluSerThrThrGlyPheThrLeutyrrGlyMet 633
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1528 CCTCCAAAAAACTGCATTCATTATTTTGNAGTAACAANAATTT---TGGTATCAGATG 1584

QY 634 LeutyrrLysProHisAspLeuGlnProGlyLysLysTyrrProthrValLeuPheIleTyr 653
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1585 ATCTTCCTCCCTCAT--TTTGATAAATCCAGAANAATATCTCTACTATTAGATGTGTAT 1641

Qy	654	GlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeu	673
Db	1642	GCAGGCCCATGTAGTCAA-----AAAGCAGACACTGTCTTCAGACTG	1683
Qy	674	Asn-----ThrLeuAlaSerLeuGlyTyrValValValVal-----IleAspAsnArg	689
Db	1684	NACTGGGCCACTTACCTTGCAGCAGACAGAAAACATTATAGTAGCTAGCTTTGATGGCAGA	1743
Qy	690	GlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIle	709
Db	1744	GGAAGTGGTTACCAAGAGATAGATCATGTCATGCAATCAACAGAGAAGCTGGGAACATTT	1803
Qy	710	GluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAsp	729
Db	1804	GAAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAATTT---TCAAAAATGGGATTTGTGGAC	1860
Qy	730	LeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyrLeuSerLeuMetAlaLeu	749
Db	1861	AAACAACGAATTTGCAATTTGGGGCTGGTCATATGGAGGGTAGTAAACCTCAATGGTCCCTG	1920
Qy	750	MetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTyrIle	769
Db	1921	GGATCGGGAAGTGGCGTGTTCAGTGTGGATAGCCGTCGCGCTGTATCCCGGTGGAG	1980
Qy	770	PheTyrAspThrGlyTyrThrGluArgTyrMetGly-----HisProAspGlnAsnGlu	787
Db	1981	TACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAGAAGACAACCTT	2040
Qy	788	GinglyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsn	807
Db	2041	GACCATTACAGAAATTCACAGCATGACGAGCAGAGCTGAAAATTTT-----AAACAAGTT	2094
Qy	808	ArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIle	827
Db	2095	GAGTACCTCCTTATTATGATGAAACAGCAGATGATTAACGTTTCACCTTTCCAGAGTCAGCTCAG	2154
Qy	828	LeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGlu	847
Db	2155	ATCTCAAGCCCTGGTCGATGTTGGATGGATTTCCAGGCAATGTGGTATACTCATGAA	2214
Qy	848	ArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyr	867
Db	2215	GACCATGGAATAGCTAGCAGCAGCAGCACCAACATATATATACCACATGAGCCACTTC	2274
Qy	868	LeuGlnGlu 870	
Db	2275	ATAAACAACAA 2283	

RESULT 7

```

US-10-501-035-34
; Sequence 34, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYMERIZATION
; TITLE OF INVENTION: ACTIVITY OF COMPOUND
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 3407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-501-035-34

Alignment Scores:
; Pred. No.: 5.4e-46
; Score: 529.00
Match

```

Percent Similarity:	42.0%	Conservative:	109
Best Local Similarity:	26.5%	Mismatches:	280
Query Match:	11.3%	Indels:	128
DB:	9	Gaps:	32
US-10-825-632-1 (1-882) x US-10-501-035-34 (1-3407)			
QY	206	ProAlaAspProAspTTPile-----AlaPheIleHis	216
Db	520	CCAAACAACACACAGTGGTCCACATGGTCCACAGTGGGTCAATAAATGGCATATGTTGG	579
QY	217	SerAsnAspIleTTPileSerAsnIleValThrArgGluGluArgLeuThrTyrVal	236
Db	580	AACATGACATTTATGTTAAATTTGAACCAAAATTTACCAAGTTTACAGAATCACATGGAGC	639
QY	237	HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal	256
Db	640	GGG-----AAAGAAGATATAATATAATGAATAACTACTGGGTT	681
QY	257	LeuGlnGluGlu---PheAspArgTyrSerGlyTyrTTPTrpCysProLysAlaGluThr	275
Db	682	TATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGTCTCCAAACGGCACTTT	741
QY	276	ThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluVal	295
Db	742	-----TTAGCATATGCCCAATTTAACGACACAGATC	774
QY	296	GluIleIleHis-----ValThrSerProMetLeuGluThrArgAlaAspSer	312
Db	775	CCACTTATTGAATACTCTTACTCTGATGAGTCACTGCAGTACCAGAAAGCTGTACGG	834
QY	313	PheArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIle	332
Db	835	GTTCATATCCAAAGGACGAGCTGTGAATCCAACTGTAAAGTCTTTGTT-----GTA	888
QY	333	MetIleAspAlaGluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPhe	352
Db	889	AATACAGACTCTCTCAGCTCAGTACCACCAATCTCCATACAAATCACTGCTCTGCT	948
QY	353	GluIleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLys	372
Db	949	TCATGTTGATA---GGGGATCACTACTTGTGTGATGTGACATGGCA-----	993
QY	373	TyrAlaTTPSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSer	392
Db	994	-----ACACAAGAAAGAAATTTCTTTGCACTGGCTC---1023	
QY	393	ProGluLeuPheIleProValGluAspAspValMetGluArgGlnArgLeuIleGluSer	412
Db	1024	-----AGGAGGATTCAAGAC	1038
QY	413	ValProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTTPileAsn	432
Db	1039	TAT-----TCGGTCTATGATATTTGTGACTATGATGATGATCCAGTGGAGATGG---AAC	1089
QY	433	IleHisAspIlePheHisValPheProGlnSerHisGluGluGluIleGluPheIlePhe	452
Db	1090	TGC-----TTAGTGGCAGCGCACACATTTGAATGAGTACTACTGGCTGG	1134
QY	453	AlaSerGluCysLysThrGlyPheArgHis-----LeuTyrLys	465
Db	1135	GTTCGGAAGATTAGCGCTTCAGAACCTCATTTTACCCTTGATGATGATAGTCTTACAG	1194
QY	466	IleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuProAlaPro	485
Db	1195	ATCATCAGC-----AATGAAGAAGGTTACAGACACTT-----1227	
QY	486	SerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTTPGluVal	505
Db	1228	TGCTATTTCCAAATAGATAAAAGACTGCACATTTTATTAACAAAGGACCTGGGAAGTC	1287
QY	506	LeuGlyArgHisGlySerAsnIleGlnValAspGluValArgGluValTyrPheGlu	525

Db 1135 TGCACCTC----- 1143
Qy 484 AlaProSerAspPheLys---CysProIleLysGluGluLeuAlaIleThrSerGlyGlu 502
Db 1144 ---CCGAAGAAAGAAAGACTGTACATTT-----ATTACAAAGAGGCC 1185
Qy 503 TrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuVal 522
Db 1186 TGGGAAGTCAATTT-----AGTATCGAAGCTCTGACCGAGATTATCTATAC 1230
Qy 523 TyrPheGluGlyThr---LysAspSerProLeuGluHisHisLeuTyrValValSerTyr 541
Db 1231 TACATAGTAAACCAATATAAGAAATGCCAGAGAGAAATCTCTATAAAATTT----- 1284
Qy 542 ValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIle 561
Db 1285 -----CAACTTACTGAC-----CACACAATGTGAAGTGCCTT 1317
Qy 562 Ser-----GlnHisCysAspPhePheLysSerLysTyrSerAsnGlnLys 576
Db 1318 AGTTGTGACCTGAATCCAGAAAGATGTCAGTATTATGCGGTATCATTTAGTAAAGAGCA 1377
Qy 577 AsnProHis-----Cys-----ValSerLeuTyrLysLeuSerSerPro 589
Db 1378 AGTACTATCAGCTGGGATGTTGGGGCCCGGCTCTGCCCTCTACACTCTACATCTAGC 1437
Qy 590 GluAspAsp-ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaG 609
Db 1438 ACGATCATAAAGAGCTGC-----GAGTCTGGAGAACATTTCTGCT-TTGATAG 1487
Qy 609 yProLeuProAspTyrThrProProGlu-----IlePheSerPheGluSerThrTh 626
Db 1488 AATGCTGCAGGATGTCAGATGCTTCAAAAAAATTTGGACTTCATTTGTTGAATCAAA 1547
Qy 626 rGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLys 646
Db 1548 AAGATTTT---TGTATCAATGATCTTGCCTCCCTCAT---TTTGATAAATCCAGAAATA 1601
Qy 646 rProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPhe 666
Db 1602 TCCTCTACTATTAGATGATATGAGGTCCTGAGTCA-----AA 1643
Qy 666 sGlyValLysTyrPheArgLeuAsn-----ThrLeuAlaSerLeuGlyTyrValVa 683
Db 1644 AGCAGATGCTTCTTCAGACTGAATGGGCCACTTACCTTCAAGTACAGAAACATCAT 1703
Qy 683 lValVal---IleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAla 702
Db 1704 AGTAGTAGCTTGTGACGGCAGAGGAAGTGGTTTACCAAGAGATAAGATCATGATGCAAT 1763
Qy 702 eLysTyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeuAl 722
Db 1764 CAACAGAGATTGGGAACACTCGAAGTTGAGATCAAAATGACACAGCCAGCAATTTGT 1823
Qy 722 aserArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGly 742
Db 1824 A---AAATGGGATTTGTGTAGTAGCAGGAGTTGCAATTTGGGGCTGTATATGGAG 1880
Qy 742 yTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAla 762
Db 1881 GTATGTAACCTCAATGGTCTCGGATCGGGAAGTGGCGTTCAGTTCGGAATAGTGTGT 1940
Qy 762 yAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGly-- 781
Db 1941 GGCACCTGTGTACGGTGGAGTACTATGACTGTGTACAGGAGGTTACATGGGTCT 2000
Qy 782 ----HisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla 800
Db 2001 CCCAATTCAGAGAGACACTTGCACCATTCAGGAACTCAACAGTCTAGCAGAGCTGA 2060
Qy 800 uLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluLeuVa 820
Db 2061 ACATTTT-----AAACAAGTTGAGTACTCTCTCTTATTCATGGAACGGCAGATGATATGT 2114

Qy 820 lHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAs 840
Db 2115 TCACCTTTCAGCAGTCAAGCTCAGATCTCAAAAGCCCTGGTGGTGTGGATTTCCA 2174
Qy 840 pLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis 860
Db 2175 AGCAATGTGGTACAGGATGAAGACACCGGATCGTAGCAGCAGCTCACCAGCACAT 2234
Qy 860 rGluLeuHisLeuLeuHisTyrLeuGlnGlu 870
Db 2235 CTATTCCACATGAGCCATTTCTCTCCAGCAG 2265

RESULT 10

US-10-932-182A-5649
; Sequence 5649, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5649
; LENGTH: 2778
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-5649

Alignment Scores:
Pred. No.: 6,89e-42 Length: 2778
Score: 490.00 Matches: 183
Percent Similarity: 41.3% Conservative: 124
Best Local Similarity: 24.6% Mismatches: 316
Query Match: 10.4% Indels: 120
Gaps: 30

US-10-825-632-1 (1-882) x US-10-932-182A-5649 (1-2778)

Qy 166 AlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGlyPheThrGlnGlnPro 185
Db 796 TCTTCCAAAGGCTATTATTGGATTAGGATTTGAGTACAGCT-----ACTATCGAACCG 849
Qy 186 LeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAspProLysLeuCys 205
Db 850 ATTCTACCACCG--GATGAAGCAGCTGATGGAATGTACATGGTTTATTGAAACTGTC 906
Qy 206 ProAla-----AspPro-----AspTrpIleAlaPheIleHisSerAsnAspIleTrp 221
Db 907 TATGCCCATTTCTCCACTGCTTACAATTACATTTATTGTTGCTAGAGAATAATTGTTT 966
Qy 222 lLysSerAsnIleValThrArgGluGluArgLeuThrTyr---ValHisAsnGluLeu 240
Db 967 CTGCAACAGGGGAATTCAGATACGCCAAGAGCTTACCCTATGATGGCTCAAAAGACATA 1026
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 1027 TTTAATGCTAAACACCAT-----TGATATACGGAAGAG 1062
Qy 261 ----PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGly 279
Db 1063 GTCTCGCGTCCGCCACCAAGCTATATGTTGG-----GCCCCAGATGAC 1104
Qy 280 GlyLysIleuArgIleLeuTyrGluAsnAspGluSerGluValGluIleHis 299
Db 1105 TCAAAAGCTGTTTTCGAAGATTTAATGACCCCGGCTTGAAGATATTCAACTAACAC 1164

Alignment Scores:

```

Pred. NO.: 6.89e-42 Length: 2778
Score: 490.00 Matches: 183
Percent Similarity: 41.3% Conservative: 124
Best Local Similarity: 24.6% Mismatches: 316
Query Match: 10.4% Indels: 120
DB: 9 Gaps: 30

US-10-825-632-1 (1-882) x US-10-932-182A-5649 (1-2778)

QY 166 AlaGlySerGlyLeuThrHisValIleAspGlyProGlnGlyPheThrGlnGlnPro 185
DB 796 TCTCCAAAGGCTATTATTGGATTAGGATTGAGTACAGT-----ACTATCGAACCG 849
QY 186 LeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAspProLysLeuCys 205
DB 850 ATTCTACACCG--GATGAGCAGCTGATGGAATGTACATGGTTATTGAAATGTCC 906
QY 206 ProAla-----AspPro-----AspTrpIleAlaPheIleHisSerAsnAspIleTrp 221
DB 907 TATGCCCATTTCTCACCTGCTTACCAATTATTTATTTGCTATGAGAATAATTTGTTT 966
QY 222 IleSerAsnIleValThrGluGluArgGluArgLeuThrTyr---ValHisAsnGluLeu 240
DB 967 CTGCAACAGGGGAATTCAGATACGCCAAGAGCTTACCCATGATGCTCAAGACATA 1026
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 1027 TTTAATGCTAAACCAT-----TGATATACGAGGAGAG 1062
QY 261 ---PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGly 279
DB 1063 GTCTCGCGTCCGACCAAGCTATATGTTG-----GCCCCAGATGAC 1104
QY 280 GlyIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHis 299
DB 1105 TCAAAAGCTGTTTTTCAGAGATTTAATGACACCGCGTGAAGATATTCACTAACCAAC 1164
QY 300 ValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGly 319
DB 1165 TATACGACAGGAGGACCATATCTTTCCAGTCTTAAATAAATACTCTAAACAGGA 1224
QY 320 ThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluArg 339
DB 1225 TCTAAAACCCCAAGTTGATTGA-----TTTAAATCAACTACAGAACGAG 1272
QY 340 IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVal 359
DB 1273 GTACTTTATTCAGTCAATACGGGT---GGTCAAAAGGATTCATTTTGTATAATGGG--- 1326
QY 360 GluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeu 379
DB 1327 -----AAATGGATTAGTCAGATTTCATTAGATTGCAAT---ACG 1365
QY 380 AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal 399
DB 1366 GACAGAAATTCGAAGGTGTTGACGTT----- 1392
QY 400 GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro 419
DB 1393 -----AAAGTATACGATGATACCGCTCTCT- 1416
QY 420 LeuIleIleTyrGluGluThrThrAspIle-----TrpIle---AsnIle 433
DB 1417 AAATCTACTTCTGTGAGAACCACTGATTTCGAATTCGTTCAATGGATGGATTGAAAGACC 1476
QY 434 HisAspIlePheHisValPheProGlnSerHisGluGluGluIleGlu-----PheIle 451
DB 1477 AAAGATATCTTAAGATTCACCAAAACCGGAATCCGAAAGACGAGTACGGATATATC 1536
QY 452 PheAlaSerGluCysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLys 471
DB 1537 GATATTCAACGGGCAATAACGGGTTCATACCTTTTATTTATTCACCAACCATCTTT--- 1593

```

```

QY 472 GluSerLysTyrLysArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysPro 491
DB 1593 ----- 1593
QY 492 IleLysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHis----- 509
DB 1594 GCCAAGATGGCATTCATTTAGCTAAAGAACTAGGAAGTACTGGGAACGGTATAATT 1653
QY 510 GlySerAsnIleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAsp 529
DB 1654 GGTTCAGAAATATCAAAACGGAT-----ACCGTATTTTATACAGCAATGAGATT 1701
QY 530 SerProLeuGluHisHisLeuTyrValValSerTyrValAsnProGlyGluValThrArg 549
DB 1702 GGAACCATGTGCACAAACACCTTTATATGATACCACTTTT---AATCTCGCAACTATAGGCAAA 1758
QY 550 Leu-----ThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
DB 1759 CTTCAAGTTATGCAAAATCCATCTGATAAATATGATTTTATGACTTTGAATTGAGCTCT 1818
QY 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSer--- 582
DB 1819 AGTGAAGATATGCCATTTTCAAAGAACTAGCCCGGTGCACCACTGATAAAGTAGCTGGA 1878
QY 583 -----LeuTyrLysLeuSerSerProGluAspAspProThr-----Cys 595
DB 1879 CCGCTTACGGAAGTACTGATCTTACCAATGCTAACGGCGATTTCGGTGTGCACTGACC 1938
QY 596 LysThrLysGluPheIleTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThr 615
DB 1939 AGGATCGAAATTTTAAAGACACATAAAGGTTTACGATCTACCAATTTACCACTTACCAAG 1998
QY 616 ProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyr 635
DB 1999 -----TCAATGGTCTTAAGTGATGTTGTAATAAATACTACATCGAAATT 2043
QY 636 LysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
DB 2044 AAACCAACCACTTAGATCCCAAGAAATAATCCGATATTAGTAAACATTTATGTTGGTGC 2103
QY 656 ProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThr 675
DB 2104 CCAGATCTCAATCATTCACCACT-----AAATCGTGTGCTTCATTCGAACAGTCTGTT 2157
QY 676 LeuAlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGly 695
DB 2158 GCTTCAGGATTGGATGCATCATCTCCAAATAGAACCAAGAGAACTGGGGGTAAAGGT 2217
QY 696 LeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnVal 715
DB 2218 TGGAGTTTCAGATCTTGGGCTCGCGGTAGATTGGGATATTGGGAGCCCTAGGATATTATT 2277
QY 716 Glu---GlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGly 734
DB 2278 GAGTCTACCAAGAAATTTTACAAAGAAACAAACCAATATAGACAAAGTCCCA 2337
QY 735 IleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeu---MetGlnArgSer 753
DB 2338 ATTTGGGTTGGTTCGTCAGCTGCTGTTGTTAGTTTGAAGCCGTAGATTGGATAATGTT 2397
QY 754 AspilePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThr 773
DB 2398 GAAACATTTCAAATACACAATGGCTGTTCACCAAGTAAACAACCTGGACACTGTATGACTCC 2457
QY 774 GlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGly 793
DB 2458 ATCTACACCGAAAGGTATATGAATCAACCATCGAAATGAAAGGGGTATTTTTGACATA 2517
QY 794 SerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHis 813
DB 2518 TCTACTATTAAAAATTACAAGTCTTTCGAATCG---CTGAACCGACTCTTTGTTATGCAC 2574

```


Db 1349 -AAAGACACTGTGGAAATGCTATTCAAATTACAGTGGCAAGTGGAGGCCATA----- 1402
Qy 508 gHisGlySerAsnIleGlnValAspGluValArgLeuValTyrPheGluGlyThrLy 528
Db 1403 -----AATATA-----TTACAGATAACACA 1422
Qy 528 sAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsnProGlyGluValTh 548
Db 1423 GGATTTCACCTGTTTATCTAGCAATGAATTTGAAGAAATACCTCGGAAGAAACACATCTA 1482
Qy 548 rArgLeuThrAspArgGlyTyrSerHisSer---CysCysIleSerGlnHis----- 564
Db 1483 CAGAAATTAGCAATTGGAAGCTATCTCCAGCAAGAAGTGTGTACTTGCCATCTAGGAA 1542
Qy 565 -----CysAspPhePheIleSerTyrSerAsnGlnLysAsnProHisCysValSe 582
Db 1543 AGAAAGGTGCCAATATTACACAGCAAGTTTCAGCGACTACGCCAAGTACTATGCATTTGT 1602
Qy 582 rLeuTyr-----LysLeuSerSerProGluAspAspP:OThrCysLysThrLy 598
Db 1603 CTGCTACGGCCAGGCATCCCATTTCCACCTTCATGATGAGCAGCACTGATCAAGAA-- 1660
Qy 598 sGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAsp----- 613
Db 1661 -----ATTAAATCTGGNAGAAACAGGAATTTGGAATGCTTTGAAAAATAT 1710
Qy 614 -TyrThrProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMe 633
Db 1711 CCAGCTGCCTAAAGAGGAATTAAGAACTTGGAAGTAGATGAATTTACTTTATGTGTACA 1770
Qy 633 tLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTy 653
Db 1771 GATGATTCCTCTCCTCAATTGACAGATCAAGAAAGTATCCCTTGCTTAATTCAGTGTGA 1830
Qy 653 tGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLe 673
Db 1831 TGGTGGTCCCTGCAGTCAGAGTGAAGTCT-----GTATTTTGGCTGT 1872
Qy 673 uAsnThrLeuAlaSerLeu-----GlyTyrValValValIleAspAsnAr 689
Db 1873 TAATTGGATATCTTATCTTGCAAGTAAGGAAGGATGGTCAATGGCTTGGTGGATGGTCG 1932
Qy 689 gGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnI 709
Db 1933 AGGAACAGCTTCCAAAGTGCACAACTCCTATGCAGTGTATCGAAAGCTGGGTGTTA 1992
Qy 709 eGluLeAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAs 729
Db 1993 TGAAGTTGAGACCCAGATTACAGCTGTCAGAAATTCATA--GAAATGGGTTTCATTGA 2049
Qy 729 pLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetAlaLe 749
Db 2050 TGAAGAAAGATAGCCATATGGGCTGCTCTATGGAGATACGTTTTCATCTGCTGCTCT 2109
Qy 749 uMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpI 769
Db 2110 TGCATCTGGAAGTGTCTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGA 2169
Qy 769 ePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro-----AspGlnAsnGl 787
Db 2170 ATATTACCGCTCTGCTACACAGAGAGATTTCATGGGTCTCCCAACAAAGAGATGATA 2229
Qy 787 uGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlyLysPheProSerGluProAs 807
Db 2230 TGAGCATTATAAGATTCACCTGTGTGGCAAGAGCAGAAATATTTCAGAAATGTAGAC-- 2287
Qy 807 nArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerI 827
Db 2288 ----TATCTTCTCATCCAGCAAGCAGAGATGATAATGTGCATTTTCAAAACTCAGCACA 2343
Qy 827 eLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnI 847
Db 2344 GATTGCTAAGCTCTGGTTAATGCACAGTGGATTTCAGGCAATGGTGTACTCTGACCA 2403

Qy 847 uArgHisSerIleArgValProGluSerGly-----GluHisTyrGluLeuHisLe 864
Db 2404 GAACCAACGGCTTA-----TCCGGCTGTCCACGAACCACTTATACACCCACAT 2451
Qy 864 uLeuHisTyrLeuGlnGlu 870
Db 2452 GACCCACTTCTTAAAGCAG 2470

RESULT 14

US-11-186-284-54
; Sequence 54, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (209)...(2491)
US-11-186-284-54

Alignment Scores:
Pred. No.: 5,63e-38 Length: 2814
Score: 454.00 Matches: 177
Percent Similarity: 41.9% Conservative: 111
Best Local Similarity: 25.8% Mismatches: 278
Query Match: 9.7% Indels: 122
DB: 14 Gaps: 27

US-10-825-632-1 (1-882) x US-11-186-284-54 (1-2814)

Qy 212 IleAlaPheIleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArg 231
Db 692 TTAGCATATGTCATCAAAACAAATATCTATTGAAACAAAGACCCAGGAGATCCACCTTTT 751
Qy 232 ArgLeuThrTyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGly 251
Db 752 CAAATTAACATTT-----AATGGAAGAGAAATAATAATTTAATGGA 793
Qy 252 ValAlaThrPheValLeuGlnGluPhe-----AspArgTyrSerGlyTyrTrp 269
Db 794 ATCCACAGACTGGGTTTATGAAGAGAAATGCTTCTCAAAATATGCT---CTCTGGTGG 850
Qy 270 CysProLysAlaGluThrThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluGlu 289
Db 851 TCTCCT-----AATGGAATAATTTTGT-----GCATATGCGGAA 883

Alignment Scores:		3.69e-36	Length:	2457
Pred. No.:		436.50	Matches:	169
Score:		37.3%	Conservative:	105
Best Local Similarity:		23.0%	Mismatches:	226
Query Match:		9.3%	Indels:	235
DB:		9	Gaps:	27
US-10-825-632-1 (1-882) x US-10-932-182A-1107 (1-2457)				
QY	212	IleAlaPheIleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArg	231	
DB	601	ATTGCATACGTCGAAGATAACAATATATATACGTCTATCTACAGAATCGAGAAACTACA	660	
QY	232	ArgLeuThr	245	
DB	661	CGGCTGTAACTGATGACGGAATCCCTCTCTCAAC	699	
QY	246	AspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluGlu---	264	PheAspArgTyr
DB	700	GGTAAGCCNGATTGGTTTACGAGAGAAGTCTTTGAGGACGAC	744	
QY	265	SerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGlyIleLeuArg	284	
DB	745	AAGCGTGTGGTGGTCACCACTGGTGAT	777	TAT
QY	285	IleLeuTyrGluIleAsnAspGluSerGluVal---	300	HisVal
DB	778	CTGGCGGTTTTAAAGATCGATGAATCAGAAGTTGGTGAGTTCATATCCCTATTATATGTC	837	
QY	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	
DB	838	CAAAATGATAAAGATGTATATCCAGAAATCGTAGATTAAAGTACCACAAAGAGTGGTACA	897	
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMethle	334	
DB	898	CCGAATCCCAT	942	GCCGAGTTGTGGTCTACTGTATGAAGACGAA
QY	335	-----AspAlaGluGlyArgIleIleAspValIleAsp	345	
DB	943	TCATCGTACCACCCTAGAGTAAATGGAGACGAAAGGTTCAAGTATATTGATTACTGAA	1002	
QY	346	LysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAlaArgAla	365	
DB	1003	CTTACATGGTGGGAGATGAGGACGCTTTAGTTAAG	1038	
QY	366	GlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArg	385	
DB	1039	-----ACTACTGATCGTTCCTCAGATATG	1062	
QY	386	LeuGlnIleValLeuIleSerProGlu	394	
DB	1063	CTAACGGTATTTTGTAGTGACACGAAATCCAAATCTTCATACGCTGTGAGAAATGAAAGT	1122	
QY	395	-----LeuPheIleProValGluAspAsp	402	
DB	1123	GCCATGGGGGATGGTGGAGATTACCATATCTCTTTATTCAGCAATGAACG	1182	
QY	403	Val---MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIle	421	
DB	1183	CTTTGTAGACCTCATAATGGTTATGTCGATATCTACCAATAAATGTTTATAATCACTTA	1242	
QY	422	IleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhePro	441	
DB	1243	GCCTACTTTGAGAATAGTAAT	1272	AACACACAC
QY	442	GlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArg	461	
DB	1272	-----	1272	
QY	462	HisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGly	481	
DB	1273	-----TACAAAGTA	1281	

QY	482	LeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGly	501	
DB	1282	-----CTGACAACTGGG	1293	
QY	502	GluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeu	521	
DB	1294	GATTGGGAGGTAGTT	1344	AATGGACCACTTGCCTTTGATTCAAAATGATCGT
QY	522	ValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyr	541	
DB	1345	CTTTATTATATTTCTACAAAGACACTCAACGAGCGCACTTCACTATATAGATTTC	1404	
QY	542	ValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIle	561	
DB	1405	CAGTCACCAACCAAGATGCATCAAGTTACTGAT	1440	ATT
QY	562	SerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLys--AsnProHisCysVa	581	
DB	1441	TTCAAGAAGGTGTATTTCGATATATCTTTCTTAGGTAGGAGGTTTCGGTTTAC	1495	
QY	581	IserLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu---	599	
DB	1496	-----TCACCTATAAAGGACCGACGTTCCATACCACAAAATTCGTT	1536	
QY	599	-----	599	
DB	1537	GATTTTCATTCCACAAAAACAGACATATATAGAAGGGAACATATTAGCGGAAACATTA	1596	
QY	600	-----PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProG	618	
DB	1597	TACTATTGGAGAGAAACGAAGAAATCAGCAAA--AATTTAGAAAGATTATTTCGATTCCCAA	1655	
QY	618	uilePheSerPheGluSerThrThrGlyPheThrLeuTyrGly-----MetLeuTyrIy	636	
DB	1656	AAAAATATTCCAGGAGTTAAATTTGGGTGAAGGATGAATCGGAGCGACATATTGGTGA	1715	
QY	636	sProHisaspLeuGlnProGly-----LysLysTyrProThrVa	649	
DB	1716	TTTCGTACGAGATCTCCCAACAGATTGTGACCAAGTTTATCGATTATTATCCAGTTT	1775	
QY	649	IleuPheIleTyrGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyVally	669	
DB	1776	TTTCTTTGCATATGGTGACCAAAATTTCTCAGCAATC-----ATCAA	1817	
QY	669	sTyrPheArgLeuAsn-----ThrLeuAlaSerLeuGlyTyrValValValVa	685	
DB	1818	GACATTTCGTAGGTTTCAATAGGTCGTAGCTTCAATTAATGCAATTTAGTTGT	1877	
QY	685	IleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLy	705	
DB	1878	CGTCGATGTCGTGTACTGGTTTCAAGGCCAAAGATTAGATCGCTGTCGCGATATA	1937	
QY	705	sMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTy	725	
DB	1938	GCITGGCGATTACGAGGCCCGGTGACCAATATATCGCGCTTCCCTTTATGTTCC---CT	1994	
QY	725	rAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyrLeuSe	745	
DB	1995	GAGTTTCGTGGACTCGCAAAAGATTTCCTTATTTGGTTGGTCGTATGGGGATATTGAC	2054	
QY	745	rLeuMetAlaLeu---MetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaPr	764	
DB	2055	TCGTGAAGACCTTGGAAAGACGACGAGGAGAACACTTTTAAGTACGGGATGCTGTGTCTC	2114	
QY	764	oValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAs	784	
DB	2115	GGTAACCGACTGGAGATTCTACGACTGCTATACCGAAAGATACATGCACACTCTCTCA	2174	
QY	784	pGlnAsnGluGlnGlyTyrTyrLeuGlySerVal-----AlaMetGlnAlaGl	800	
DB	2175	AGAAAACCTTTCGCGGTGCTGGAATCAAGTGTCCACAATGTCACAGCGTTG-----	2226	

us-10-825-632-1.p2n.rnpbn

Tue Apr 18 08:18:33 2006

```

Qy      800 uLysPheProSerGluProAsnAtcLeuLeuLeuHisGlyPheLeuAspGluAsnVa 820
Db      2227 -----GCCCAAGTAAACAGGTTCTACTAATGCGTGTACAGGGGATGCAATGT 2276
Qy      820 lHisPheAlaHisThrSerIleLeuLeuSerPheLeu---ValArgAlaGlyLysProTy 839
Db      2277 TCACITCCAAAATTCTCTGAATTTCTAGACCTTTTGAACCTGAACAATGTCGAAAATT 2336
Qy      839 rAspLeuGlnIleTyProGlnGluArgHisSerIleArg 852
Db      2337 CGATGTTTCATGTATTCCTGTATTCGACACAGTATTAAGA 2376

```

Search completed: April 15, 2006, 01:49:57
Job time : 930 secs

CTGTTTCATGTATTCCTGTATTCGACACAGTATTAAGA

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 19:35:26 ; Search time 15351 Seconds
(without alignments)
11553.095 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgctaaagcctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sy.*
12: gb.un.*
13: gb.vi.*
14: gb.htg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3120	100.0	3120	6 AR651452	AR651452 Sequence
2	3120	100.0	3127	8 AF221634	AF221634 Homo sapi
3	3106.4	99.6	4829	6 AR631283	AR631283 Sequence
4	3106.4	99.6	4829	6 AR608735	AR608735 Sequence
5	3105.2	99.5	3143	6 AX354793	AX354793 Sequence
6	3101.8	99.4	3106	6 AX342633	AX342633 Sequence
7	2915.2	93.4	4535	8 BC040203	BC040203 Homo sapi
8	2824.4	90.5	4685	6 AR631288	AR631288 Sequence
9	2824.4	90.5	4685	6 AR608745	AR608745 Sequence
10	2806.4	89.9	4676	6 AR631287	AR631287 Sequence
11	2806.4	89.9	4676	6 AR608743	AR608743 Sequence
12	2805.4	89.9	3030	8 AY354202	AY354202 Homo sapi
13	2759.4	88.4	2797	6 AR448400	AR448400 Sequence
14	2656.4	85.1	3125	8 BC030688	BC030688 Homo sapi
15	2645.8	84.8	2649	6 AX354795	AX354795 Sequence
16	2644.4	84.8	2656	8 AY172659	AY172659 Homo sapi
17	2644.4	84.8	2671	6 AR631278	AR631278 Sequence
18	2644.4	84.8	2671	6 AX608725	AX608725 Sequence

19	2500.4	80.1	4523	6 AR631281	AR631281 Sequence
20	2500.4	80.1	4523	6 AR608731	AR608731 Sequence
21	2490	79.8	2842	6 AX405770	AX405770 Sequence
22	2445.4	78.4	2778	8 AK000290	AK000290 Homo sapi
23	2273.4	72.9	4799	9 BC043124	BC043124 Mus muscu
24	2273.4	72.9	4799	9 BC059222	BC059222 Mus muscu
25	2208	70.8	2510	6 AX338497	AX338497 Sequence
26	2132	68.3	2668	6 AX405771	AX405771 Sequence
27	2056.4	65.9	4309	6 AR631284	AR631284 Sequence
28	2056.4	65.9	4309	6 AR608737	AR608737 Sequence
29	1994.8	63.9	2161	6 BD157001	BD157001 Primer fo
30	1994.8	63.9	2161	6 AX878058	AX878058 Sequence
31	1994.8	63.9	2161	8 AK027826	AK027826 Homo sapi
32	1369	43.9	18341	14 AC161443	AC161443 Mus muscu
33	1347	43.2	1669	6 AR651454	AR651454 Sequence
34	1347	43.2	1669	8 AF221636	AF221636 Homo sapi
35	1330.8	42.7	2098	6 CQ721272	CQ721272 Sequence
36	1036	33.2	1197	6 AR651453	AR651453 Sequence
37	1036	33.2	1197	8 AF221635	AF221635 Homo sapi
c	967.4	31.0	194188	9 AC163387	AC163387 Mus muscu
39	926.4	29.7	1356	6 AR631282	AR631282 Sequence
40	926.4	29.7	1356	6 AX608733	AX608733 Sequence
41	901.6	28.9	3383	5 BC071112	BC071112 Xenopus l
42	882.2	28.3	3399	9 BC057631	BC057631 Mus muscu
43	862.6	27.6	2660	6 AX921883	AX921883 Sequence
44	862.6	27.6	2660	6 AX921885	AX921885 Sequence
45	862.2	27.6	1449	5 CR353491	CR353491 Gallus ga

ALIGNMENTS

RESULT 1
AR651452
LOCUS AR651452 3120 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence, 2 from patent US 6881564.
ACCESSION AR651452
VERSION AR651452.1 GI:62795938
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3120)
AUTHORS Abbott, C.A. and Gorrell, M.D.
TITLE Dipeptidyl peptidases
JOURNAL Patent: US 6881564-A 2 19-APR-2005;
The University of Sydney, Sydney;
AUX;
FEATURES
source Location/Qualifiers
1..3120
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN	Query Match	Best Local Similarity	Score	DB 6;	Length	3120;
	100.0%;	100.0%;	100.0%;	0;	0;	0;
	Mismatches	0;	Mismatches	0;	Indels	Gaps
Qy	1 AAGTGCTAAAGCCTCCGAGCCCAAGCGCTGCTACTCGCCGCTGCTCTTCTTAGTCCG	60				
Db	1 AAGTGCTAAAGCCTCCGAGCCCAAGCGCTGCTACTCGCCGCTGCTCTTCTTAGTCCG	60				
Qy	61 CGTTTCGCGCTGGTGTGTTCACCGCGCGCGCGCGCGAGCAAGCCACTGCAACACGAGACCG	120				
Db	61 CGTTTCGCGCTGGTGTGTTCACCGCGCGCGCGCGAGCAAGCCACTGCAACACGAGACCG	120				
Qy	121 GAGTGGAGGCGCGCAGCATGAAGCGCGCGCGCTTCCATAGCGCACCTCGGGACGG	180				
Db	121 GAGTGGAGGCGCGCAGCATGAAGCGCGCGCGCTTCCATAGCGCACCTCGGGACGG	180				
Qy	181 TCCGGCGCGCGCGCGGGGAGGAAATTCACATGCGCAGCGCAATCGAACAACAGACAG	240				
Db	181 TCCGGCGCGCGCGCGGGGAGGAAATTCACATGCGCAGCGCAATCGAACAACAGACAG	240				

Db	2401	TTAGATCGTGTGGGATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTA	2460
Qy	2461	ATGCAGAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATC	2520
Db	2461	ATGCAGAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATC	2520
Qy	2521	TTCTATGATACAGGATACACGGAAGTATATATGGTCACTGACACAGAAATGAACAGGGC	2580
Db	2521	TTCTATGATACAGGATACACGGAAGTATATATGGTCACTGACACAGAAATGAACAGGGC	2580
Qy	2581	TATTACTTAGAGTCTGTGGCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTA	2640
Db	2581	TATTACTTAGAGTCTGTGGCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTA	2640
Qy	2641	CTGCTCTTACATGCTTCTCGGATGAGATGCAATTTTGCACATACACAGTATATTACTG	2700
Db	2641	CTGCTCTTACATGCTTCTCGGATGAGATGCAATTTTGCACATACACAGTATATTACTG	2700
Qy	2701	AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC	2760
Db	2701	AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC	2760
Qy	2761	AGCATAGAGTTCCTGAATCGGAGAAATATGAACTGCAATCTTTTGGCACTACCTTCAA	2820
Db	2761	AGCATAGAGTTCCTGAATCGGAGAAATATGAACTGCAATCTTTTGGCACTACCTTCAA	2820
Qy	2821	GAACACCTTGATCAGTATGCTGCTTAAAGTGATATAATTTGACCTGTGAGAAC	2880
Db	2821	GAACACCTTGATCAGTATGCTGCTTAAAGTGATATAATTTGACCTGTGAGAAC	2880
Qy	2881	TCTCTGTATACACTGGTATTTAAACCAATCAGAGAGTTTAAATCAACAGAAACACAGA	2940
Db	2881	TCTCTGTATACACTGGTATTTAAACCAATCAGAGAGTTTAAATCAACAGAAACACAGA	2940
Qy	2941	ATTGATCATCACAATTTTGATPACTGCTGATGTAACATCTACTCTGAAATAAATGTGGTG	3000
Db	2941	ATTGATCATCACAATTTTGATPACTGCTGATGTAACATCTACTCTGAAATAAATGTGGTG	3000
Qy	3001	CCATGAGGGGTACAGTTTGGTGTAGTAATCTAATACCTTAACCCACACATGCTCAAAA	3060
Db	3001	CCATGAGGGGTACAGTTTGGTGTAGTAATCTAATACCTTAACCCACACATGCTCAAAA	3060
Qy	3061	TCAAATGATACATATCTCTGAGAGCCAGCAATACCAATGAGATTTACTTAAAAAAAAA	3120
Db	3061	TCAAATGATACATATCTCTGAGAGCCAGCAATACCAATGAGATTTACTTAAAAAAAAA	3120
RESULT 2			
AF221634			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

FEATURES		Location/Qualifiers	
source		1..3127	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/chromosome="15"	
		/map="15q22"	
gene		1..3127	
		/gene="DPP8"	
CDS		214..2862	
		/gene="DPP8"	
		/note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; alternatively spliced forms of this mRNA are found in AF221635, AF221636 and AF221637"	
		/codon_start=1	
		/product="dipeptidyl peptidase 8"	
		/protein_id="AAG29766.1"	
		/db_xref="GI:11095188"	
		/translation="MAAAWETEQLGVEIFETADCEENIESQDRPKLEPFFVERYSWQ LKLLADTRKHGYMMAKAPHDPMFKNDPDGPHSDRIYYLAMSGENRENTLIFYSEI PKTINRAALVMSLWKLPLDLOATLDYGMYSRBEELRKRKRGITGYASDYHQSSG TFLFOAGSGIYHVKGDPGFTQOPLRPNLVETSCPNIMRMDPKLCPADPMWIAFHSN DIWISNIVTREERLTYVHNELANMEEDARSAGVATVLOEEFDRVSGVWMCPEKATT PSQKILRLIYBEENDESEVEIIVHTVSPLMETRADSFYRPTGTANPKVTPKMSIIMI DAEGRIDVIDKELIOPPEILFEGVEYIARAGWTPEGKYAWSILLDRSQRLQLVLIS PELFIPVEDVNMERQLIESVPDSVTPLIYYEETDIWNIHDIHFVFPQSHBEEIEF IPASECKTFRHLKYITSLKESKYKRSGLPAPSPDKPCPKIEETAITSGEWEVLGR HGSNIQDVSRVRLVYFEGTKDPLRHLHYVVSYPGEVTLRLTRDGYSHSCISQHD PFISKYSNKNCHCVSLYKLSPEDDPTCKTKEFWATILDSAGPLDPTPTPEIAPSPES TTGTUUGMLYKPHDLQPKKPTVILFIYDVGQVQLVNNRFPKGVKVFRLNTLASGVY VVVIDNRSGHRLGPEFAGKYMGOIEIDDOVGLQLALASRYDFIDLRDRLHWSY GGYLSLMALMQRSDI PRVAIAGAPVTLMIFYDTGYTERYMGHPDQNGQYGLGSVAMQ AEKFPSEPNRLLLHGFLDENVHFAHTSILLSFLVRAGKPYDLQIYPOERHSIRVPES GEHYELHLLHYLQENLGSRIAALKVI"	
ORIGIN		Query Match 100.0%; Score 3120; DB 8; Length 3127;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	AAGTGTAAAGCCTCCGAGGCCAAGCCCGCTCTACTCGCCGCGCTGCTTCTTAGTCCG	60
Db	1	AAGTGTAAAGCCTCCGAGGCCAAGCCCGCTCTACTCGCCGCGCTGCTTCTTAGTCCG	60
Qy	61	CGTTCCGCGCTCGGTTGTACCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAG	120
Db	61	CGTTCCGCGCTCGGTTGTACCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAG	120
Qy	121	GAGTGGAGCGCGCGCGAGCATGAAGCGCGCGCGCGCTTCCATAGCGACGTCGCGG	180
Db	121	GAGTGGAGCGCGCGCGAGCATGAAGCGCGCGCGCGCTTCCATAGCGACGTCGCGG	180
Qy	181	TCGCGGCGCGCGCGCGGAGAAATGCAACATGCGAGCGAGAGAGAGAGAGAGAG	240
Db	181	TCGCGGCGCGCGCGCGGAGAAATGCAACATGCGAGCGAGAGAGAGAGAGAGAG	240
Qy	241	CTGGGCTTTCAGATATTGGAACCTGCGGAGCTGTGAGGAGAAATATTGAATCA	300
Db	241	CTGGGCTTTCAGATATTGGAACCTGCGGAGCTGTGAGGAGAAATATTGAATCA	300
Qy	301	CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCTCGAGTCACTTTAAAGCTG	360
Db	301	CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCTCGAGTCACTTTAAAGCTG	360
Qy	361	CCCGATACCAAGAAATATCATGCTACATGATGGTGAAGGACCAACATGATTT	420
Db	361	CCCGATACCAAGAAATATCATGCTACATGATGGTGAAGGACCAACATGATTT	420
Qy	421	GTGAAGAGGATGATCCAGATGACCTCATTCAGACAGATCTATTACCTTGCCAT	480
Db	421	GTGAAGAGGATGATCCAGATGACCTCATTCAGACAGATCTATTACCTTGCCAT	480

QY	481	GGTGAGACAGAGAAAATACACTGTTTATTTCTGAAATTC	Db	1561	TTTTTTTTGGCTCTGAATGC	1620
Db	481	GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTC	QY	1621	TTAAAGGAACAAATATATAACGATCCAGTGGTGGCTGCCTC	1680
QY	541	GCAGTCTTAATAGCTCTCTGTGAGAGCCCTCTTTTGGATCT	Db	1621	TTAAAGGAACAAATATATAACGATCCAGTGGTGGCTGCCTC	1680
Db	541	GCAGTCTTAATAGCTCTCTGTGAGAGCCCTCTTTTGGATCT	QY	1681	TGTCCTATCAAAAGAGGATAGCAATTTACAGTGGTGAATGG	1740
QY	601	GGAAATGTTATCTCGAGAGAGAACTATTAAGAGAGAAAG	Db	1681	TGTCCTATCAAAAGAGGATAGCAATTTACAGTGGTGAATGG	1740
Db	601	GGAAATGTTATCTCGAGAGAGAACTATTAAGAGAGAAAG	QY	1741	GGATCTAATATCAAAAGTGTGATGAAGCTGGTATATTTTGA	1800
QY	661	ATTGCTCTTACGATTATCAACAAGAGAGTGGAAACATTT	Db	1741	GGATCTAATATCAAAAGTGTGATGAAGCTGGTATATTTTGA	1800
Db	661	ATTGCTCTTACGATTATCAACAAGAGAGTGGAAACATTT	QY	1801	TCCTCTTTAGAGCATCACTGTACGTAGTCAGTTAACCTGG	1860
QY	721	ATTATACGTAAGATGGAGGCCACAGAGTTTACGCAACAA	Db	1801	TCCTCTTTAGAGCATCACTGTACGTAGTCAGTTAACCTGG	1860
Db	721	ATTATACGTAAGATGGAGGCCACAGAGTTTACGCAACAA	QY	1861	CTGACTGACCGTGGCTACTCACATTTCTGTGTCAGTCAG	1920
QY	781	CTAGTGGAACTAGTTGTCCTCAACATACGGAATGGATCC	Db	1861	CTGACTGACCGTGGCTACTCACATTTCTGTGTCAGTCAG	1920
Db	781	CTAGTGGAACTAGTTGTCCTCAACATACGGAATGGATCC	QY	1921	AGTAAAGTATAGTAAACCAAGAAATCCACACTGTGTGTC	1980
QY	841	GACTGGATGCTTTTATACATAGCAACGATATTTTGGATAT	Db	1921	AGTAAAGTATAGTAAACCAAGAAATCCACACTGTGTGTC	1980
Db	841	GACTGGATGCTTTTATACATAGCAACGATATTTTGGATAT	QY	1981	GAAGATGACCCAACTTGCAAAACAAAGGAAATTTTGGGCC	2040
QY	901	GAAAGAGACTCACTTATGTGCACAAATGAGCTAGCCAACT	Db	1981	GAAGATGACCCAACTTGCAAAACAAAGGAAATTTTGGGCC	2040
Db	901	GAAAGAGACTCACTTATGTGCACAAATGAGCTAGCCAACT	QY	2041	CCCTCTCTGACTATACCTCTCCAGAAATTTCTCTTTGAA	2100
QY	961	GCTGGAGTGGCTACCTTTGTTCTCCAAAGAAATTTGATAG	Db	2041	CCCTCTCTGACTATACCTCTCCAGAAATTTCTCTTTGAA	2100
Db	961	GCTGGAGTGGCTACCTTTGTTCTCCAAAGAAATTTGATAG	QY	2101	TTGTATGGATGCTCTACAAGCTCTAGTCTACAGCTCGAAG	2160
QY	1021	TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCT	Db	2101	TTGTATGGATGCTCTACAAGCTCTAGTCTACAGCTCGAAG	2160
Db	1021	TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCT	QY	2161	CTGTTTCATATATGTGGTCTCTCAGTGCAAGTGTGTGAAT	2220
QY	1081	AATGATGAATCTGAGTGGAAATTTATCATGTTACATCCCT	Db	2161	CTGTTTCATATATGTGGTCTCTCAGTGCAAGTGTGTGAAT	2220
Db	1081	AATGATGAATCTGAGTGGAAATTTATCATGTTACATCCCT	QY	2221	TATTTCCGCTTGAATACCTAGCTCTCTAGCTTATGTGGTGT	2280
QY	1141	GCAGATCAATCCGTTATCTTAAACAGGTACAGCAATCTTA	Db	2221	TATTTCCGCTTGAATACCTAGCTCTCTAGCTTATGTGGTGT	2280
Db	1141	GCAGATCAATCCGTTATCTTAAACAGGTACAGCAATCTTA	QY	2281	GGATCCTGTCAACCGAGGGCTTAAATTTGAAGCGCCTTTA	2340
QY	1201	TCAGAAATATATGATGCTGAGAGAGGATCATAGTGTCTAG	Db	2281	GGATCCTGTCAACCGAGGGCTTAAATTTGAAGCGCCTTTA	2340
Db	1201	TCAGAAATATATGATGCTGAGAGAGGATCATAGTGTCTAG	QY	2341	GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT	2400
QY	1261	CAACCTTTGAGATTCATTTTGAAGAGTTGAATATTTGCC	Db	2341	GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT	2400
Db	1261	CAACCTTTGAGATTCATTTTGAAGAGTTGAATATTTGCC	QY	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGAT	2460
QY	1321	GAGGAAATATGCTTTGGTCCATCTCTACATGATCGCTCC	Db	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGAT	2460
Db	1321	GAGGAAATATGCTTTGGTCCATCTCTACATGATCGCTCC	QY	2461	ATCGAGAGTTCAGATATCTTCAAGGTTGCTATTTGCTGGG	2520
QY	1381	TTGATCTCACCTGAAATTTATTTCCAGTGAAGATGATGTT	Db	2461	ATCGAGAGTTCAGATATCTTCAAGGTTGCTATTTGCTGGG	2520
Db	1381	TTGATCTCACCTGAAATTTATTTCCAGTGAAGATGATGTT	QY	2521	TTCTATGATACAGGATACACGGAACTGTTATATGGGTGCT	2580
QY	1441	ATTGAGTCAGTGGCTGATTCGTGAGCGCACTTAAATTTAT	Db	2521	TTCTATGATACAGGATACACGGAACTGTTATATGGGTGCT	2580
Db	1441	ATTGAGTCAGTGGCTGATTCGTGAGCGCACTTAAATTTAT	QY	2581	TATTAATTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCC	2640
QY	1501	TGGATAAATATCATGACATCTTTCAATGTTTTTCCCAAAGT	Db	2581	TATTAATTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCC	2640
Db	1501	TGGATAAATATCATGACATCTTTCAATGTTTTTCCCAAAGT	QY	2641	CTGCTCTTACATGCTTCTTGGATGGAATGTCATTTTGGACAT	2700
QY	1561	TTTTTTTTTGGCTCTGAATGC				

Db	2641	CTGCTCTTACATGGTTTCCTGGATGAGAAATGTCATTTTGCACATACCATATATTACTG	2700
Qy	2701	AGTTTTTTTAGTGAAGGCTGGAAGCCATATGATTTTACAGATCTATCTCCAGGAGACAC	2760
Db	2701	AGTTTTTTTAGTGAAGGCTGGAAGCCATATGATTTTACAGATCTATCTCCAGGAGACAC	2760
Qy	2761	AGCATAGAGTTCTCGAATCGGGAGAAATTATGAACTGCAATCTTTTGGACTACCTTCAA	2820
Db	2761	AGCATAGAGTTCTCGAATCGGGAGAAATTATGAACTGCAATCTTTTGGACTACCTTCAA	2820
Qy	2821	GAATACCTTGGATCAGTATGCTGCTCTAAAGTGATATATTTTGACCTGTGTAGAAC	2880
Db	2821	GAATACCTTGGATCAGTATGCTGCTCTAAAGTGATATATTTTGACCTGTGTAGAAC	2880
Qy	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA	2940
Db	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA	2940
Qy	2941	ATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTGGAATAAATGTGGTG	3000
Db	2941	ATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTGGAATAAATGTGGTG	3000
Qy	3001	CAATGAGGGTCTACGGTTTGTGTAGTAACTTAATACCTTAACCCCAATGCTCAAAA	3060
Db	3001	CAATGAGGGTCTACGGTTTGTGTAGTAACTTAATACCTTAACCCCAATGCTCAAAA	3060
Qy	3061	TCAAATGATACATATTTCTCGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAA	3120
Db	3061	TCAAATGATACATATTTCTCGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAA	3120
RESULT 3			
AR631283			
LOCUS			
DEFINITION			
SEQUENCE 12 from patent US 6844180.			
ACCESSION			
AR631283			
VERSION			
AR631283.1			
KEYWORDS			
GI:59770927			
SOURCE			
Unknown.			
ORGANISM			
Unknown.			
Unclassified.			
REFERENCE			
1 (bases 1 to 4829)			
AUTHORS			
Oj, S., Akimanya, K.O., Riviere, P.J.M. and Junien, J.-L.			
TITLE			
Serine protease genes related to DpPiv			
JOURNAL			
Patent: US 6844180-A 12 18-JAN-2005;			
Ferring BV;;			
NLX;			
FEATURES			
Location/Qualifiers			
source			
1..4829			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Query Match			
Best Local Similarity 99.6%; Score 3106.4; DB 6; Length 4829;			
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;			
Qy	1	AAGTGTAAAGCCTCCGAGGCCAAGCCCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCG	60
Db	1	AAGTGTAAAGCCTCCGAGGCCAAGCCCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCG	60
Qy	61	CGTTCCGCGCTGGTGTGTCACCGCGCGCGCGCGAGAGAGCCACTGCAACAGGACCG	120
Db	61	CGTTCCGCGCTGGTGTGTCACCGCGCGCGCGCGAGAGAGCCACTGCAACAGGACCG	120
Qy	121	GAGTGAAGGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCATGTCGGGACG	180
Db	121	GAGTGAAGGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCATGTCGGGACG	180
Qy	181	TCCGGGCGGGCGGGGGAGGAAATGCAATGCAATGCGCGCGCGCGAGTGAACAGACAG	240
Db	181	TCCGGGCGGGCGGGGGAGGAAATGCAATGCAATGCGCGCGCGAGTGAACAGACAG	240

Db	1321	GAGGAAATATGCTTGGTGCATCCTCTACAGTTCGCTCCAGACTCGCTCAGAGTAGTG	1380
Qy	1381	TTGATCTCACCTGAATATTATTTATCCAGTAGAAGATGATGTTTATGGAAAGCGAGACTC	1440
Db	1381	TTGATCTCACCTGAATATTATTTATCCAGTAGAAGATGATGTTTATGGAAAGCGAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCCTGATTTCTGTGACGCCACTAAATTTATCTATGAAGAACAACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCCTGATTTCTGTGACGCCACTAAATTTATCTATGAAGAACAACAGACATC	1500
Qy	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTACAGAGAGGAAATTGAG	1560
Db	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTACAGAGAGGAAATTGAG	1560
Qy	1561	TTTATTTTTCCTCTGATGCATCAACAGAGTTTCCGTCATTATACAAAATTACATCTATT	1620
Db	1561	TTTATTTTTCCTCTGATGCATCAACAGAGTTTCCGTCATTATACAAAATTACATCTATT	1620
Qy	1621	TTTAAAGGAAGCAATATTAACGATCCAGTGGTGGCTGCTCCAAAGTCAAGTCAATCAAG	1680
Db	1621	TTTAAAGGAAGCAATATTAACGATCCAGTGGTGGCTGCTCCAAAGTCAAGTCAATCAAG	1680
Qy	1681	TGTCCTATCAAGAGGAGATAGCAATTAACAGTGGTGGATGGGAAGTTCTTGGCGGCAT	1740
Db	1681	TGTCCTATCAAGAGGAGATAGCAATTAACAGTGGTGGATGGGAAGTTCTTGGCGGCAT	1740
Qy	1741	GGATCTAATATCAAGTGGTGAAGTGAAGTGAAGTGGTGAATTTTGAAGGCAACAAAGAC	1800
Db	1741	GGATCTAATATCAAGTGGTGAAGTGAAGTGGTGAATTTTGAAGGCAACAAAGAC	1800
Qy	1801	TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGCAAGG	1860
Db	1801	TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGCAAGG	1860
Qy	1861	CTGACTGACGCTGGCTACTCACATTTCTGTGTCATCAGTCAGCACTGTGACTTTCTTTATA	1920
Db	1861	CTGACTGACGCTGGCTACTCACATTTCTGTGTCATCAGTCAGCACTGTGACTTTCTTTATA	1920
Qy	1921	AGTAAAGTAAAGACAGAGAAATCCACACTGTGTGCTCTTTTGAAGGCAACAAAGAC	1980
Db	1921	AGTAAAGTAAAGACAGAGAAATCCACACTGTGTGCTCTTTTGAAGGCAACAAAGAC	1980
Qy	1981	GAAGATGACCCAACTTGCAGAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTTGCAGAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACA	2100
Db	2041	CCTCTTCTGACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACA	2100
Qy	2101	TTGTATGGATGCTCTACAGCCTCATGATCAGCCTGGAAGAAATATCTCTATCTGTG	2160
Db	2101	TTGTATGGATGCTCTACAGCCTCATGATCAGCCTGGAAGAAATATCTCTATCTGTG	2160
Qy	2161	CTGTTTCATATATGGTGG--TCCTCAGGTGCAGTTGGTGAATAATTCGGTTTAAAGGAGTCA	2218
Db	2161	CTGTTTCATATATGGTGGTCTCCTCAGGTGCAGTTGGTGAATAATTCGGTTTAAAGGAGTCA	2220
Qy	2219	AGTATTTTCGCTTGAATACCTTAGCCTCTCTAGGTGTTAGTGGTTGTAGTGATAGACAACA	2278
Db	2221	AGTATTTTCGCTTGAATACCTTAGCCTCTCTAGGTGTTAGTGGTTGTAGTGATAGACAACA	2280
Qy	2279	GGGATCTCTGTACCCAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAA	2338
Db	2281	GGGATCTCTGTACCCAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAA	2340
Qy	2339	TAGNAATTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTCAATTG	2398
Db	2341	TAGNAATTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTCAATTG	2400
Qy	2399	ACTTAGATCTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCTGATGGCAT	2458
Db	2401	ACTTAGATCTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCTGATGGCAT	2460
Qy	2459	TAATGTCAGAGTCAAGATCTTTCAGGTTGCTATTCCTGGGGCCCGAGTCACTCTGTGGA	2518
Db	2461	TAATGTCAGAGTCAAGATCTTTCAGGTTGCTATTCCTGGGGCCCGAGTCACTCTGTGGA	2520
Qy	2519	TCCTTCATGATACAGGATACAGGAGAGTTATATGGTCACTCCCTGACCAAGATGAACAGG	2578
Db	2521	TCCTTCATGATACAGGATACAGGAGAGTTATATGGTCACTCCCTGACCAAGATGAACAGG	2580
Qy	2579	GCTATTTCTTAGATCTGTGGCCATGACAGAGAAAGTCCCTCTGACCAAAATCGTT	2638
Db	2581	GCTATTTCTTAGATCTGTGGCCATGACAGAGAAAGTCCCTCTGACCAAAATCGTT	2640
Qy	2639	TACTCTCTTATCATGTTTCTCGATGAGAAATGTCCATTTTGGACATACAGTATATTAC	2698
Db	2641	TACTCTCTTATCATGTTTCTCGATGAGAAATGTCCATTTTGGACATACAGTATATTAC	2700
Qy	2699	TGAGTTTTTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTATTCCTCAGGAGAGAC	2758
Db	2701	TGAGTTTTTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTATTCCTCAGGAGAGAC	2760
Qy	2759	ACAGCATAAAGATTCCTGAAATCGGGAGAACATTTATGAACTGCACTCTTTTGCACTACCTTC	2818
Db	2761	ACAGCATAAAGATTCCTGAAATCGGGAGAACATTTATGAACTGCACTCTTTTGCACTACCTTC	2820
Qy	2819	AGAAACCTTGGATCAGTATTCCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGA	2878
Db	2821	AGAAACCTTGGATCAGTATTCCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGA	2880
Qy	2879	ACTCTCTGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACA	2938
Db	2881	ACTCTCTGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACA	2940
Qy	2939	GAATTGATCATCAGATTTTGTATCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG	2998
Db	2941	GAATTGATCATCAGATTTTGTATCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG	3000
Qy	2999	TGCCATGCAAGGCTCTACGTTTGTGTAGTAAATCTAATACCTTAACCCACACACACACAC	3058
Db	3001	TGCCATGCAAGGCTCTACGTTTGTGTAGTAAATCTAATACCTTAACCCACACACACACAC	3060
Qy	3059	AATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCATAGAAATTTACTTAAARAAA	3118
Db	3061	AATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCATAGAAATTTACTTAAARAAA	3120
Qy	3119	AA 3120	
Db	3121	AA 3122	
RESULT 4			
AX608735			
LOCUS			
Sequence 12 from Patent WO0231134.			
DEFINITION			
AX608735			
ACCESSION			
AX608735.1 GI:28404303			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1 .4829			
QI.S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.			
Novel serine protease genes related to dppiv			
Patent: WO 0231134-A 12 18-APR-2002;			
Ferring BV (NL)			
FEATURES			
source			
Location/Qualifiers			
1..4829			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			

AX608735 4829 bp DNA linear PAT 17-FEB-2003

Sequence 12 from Patent WO0231134.

AX608735

AX608735.1 GI:28404303

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1 .4829

QI.S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.

Novel serine protease genes related to dppiv

Patent: WO 0231134-A 12 18-APR-2002;

Ferring BV (NL)

FEATURES

source

Location/Qualifiers

1..4829

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

QY	2101	TTCTATGGGATGCTTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG	2160
Db	2101	TTGTATGGGATGCTTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG	2160
QY	2161	CTGTTTCATATGTTGG--TCCTCAGGTGAGTGTGTAATATCGGTTTAAAGAGTCA	2218
Db	2161	CTGTTTCATATGTTGGTCTCCCTCAGGTGAGTGTGTAATATCGGTTTAAAGAGTCA	2220
QY	2219	AGTATTTCCGTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTGATAGACAACA	2278
Db	2221	AGTATTTCCGTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTGATAGACAACA	2280
QY	2279	GGGATCTCTGTCACCGAGGCTTAAATTTGAAGGCCCTTTAAATATAAAATGGGTCAAA	2338
Db	2281	GGGATCTCTGTCACCGAGGCTTAAATTTGAAGGCCCTTTAAATATAAAATGGGTCAAA	2340
QY	2339	TAGAAATTTGAGGATCAGGTGGAAGGATCCCAATATCTAGCTCTCTCGATATGATTTCAATG	2398
Db	2341	TAGAAATTTGAGGATCAGGTGGAAGGATCCCAATATCTAGCTCTCTCGATATGATTTCAATG	2400
QY	2399	ACTTAGATCGTGTGGGCATCCAGCGCTGGTCTATGGAGGATACCTCTCCCTGATGGCAT	2458
Db	2401	ACTTAGATCGTGTGGGCATCCAGCGCTGGTCTATGGAGGATACCTCTCCCTGATGGCAT	2460
QY	2459	TAATGCAGAGTTCAGATATCTTACAGGCTTCTATGCTGGGGCCCCAGTCACTCTGTGGA	2518
Db	2461	TAATGCAGAGTTCAGATATCTTACAGGCTTCTATGCTGGGGCCCCAGTCACTCTGTGGA	2520
QY	2519	TCTTCTATGATACAGGATACACGGAAGTCTATATGGTCACTGACCAAGATGAACAGG	2578
Db	2521	TCTTCTATGATACAGGATACACGGAAGTCTATATGGTCACTGACCAAGATGAACAGG	2580
QY	2579	GCTATTACTTAGGATCTGTGGCCTGATGAGGAGGATGCTGCTGACCAAGATGCTT	2638
Db	2581	GCTATTACTTAGGATCTGTGGCCTGATGAGGAGGATGCTGCTGACCAAGATGCTT	2640
QY	2639	TACTGCTCTTACATGTTTCTGATGAGATGTCATTTTGCACATACAGTATATTAC	2698
Db	2641	TACTGCTCTTACATGTTTCTGATGAGATGTCATTTTGCACATACAGTATATTAC	2700
QY	2699	TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAGAC	2758
Db	2701	TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAGAC	2760
QY	2759	ACAGCATAGAGTTCTCGAATCGGGAGAACATTATGAACCTGACATCTTTTGCACTACCTTC	2818
Db	2761	ACAGCATAGAGTTCTCGAATCGGGAGAACATTATGAACCTGACATCTTTTGCACTACCTTC	2820
QY	2819	AAGAAAACTTTGGATCAGTATGCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGA	2878
Db	2821	AAGAAAACTTTGGATCAGTATGCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGA	2880
QY	2879	ACTCTCTGTATACACTGGCTATTTAACCAATGAGGAGGTTTAAATCAACAGAAAAACA	2938
Db	2881	ACTCTCTGTATACACTGGCTATTTAACCAATGAGGAGGTTTAAATCAACAGAAAAACA	2940
QY	2939	GAATTGATCATCACAATTTTGATACCTGCCATGTAACATCTACTCTCTGAAATAAATGTGG	2998
Db	2941	GAATTGATCATCACAATTTTGATACCTGCCATGTAACATCTACTCTCTGAAATAAATGTGG	3000
QY	2999	TGCCATGAGGGGCTCAGGTTTGGTGTAGTAACTTAATACCTTAACCCCACTGCTCAA	3058
Db	3001	TGCCATGAGGGGCTCAGGTTTGGTGTAGTAACTTAATACCTTAACCCCACTGCTCAA	3060
QY	3059	AATCAAAATGATATATTTCTGAGAGCCAGCAATACCAAGTAATTAATAAAAAAA	3118
Db	3061	AATCAAAATGATATATTTCTGAGAGCCAGCAATACCAAGTAATTAATAAAAAAA	3120
QY	3119	AA 3120	
Db	3121	AA 3122	

AX354793
Sequence 1 from Patent WO0179473.
AX354793
AX354793.1 GI:18619526
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

Meyers, R.A. and Williamson, M.
21953, a human prolyl oligopeptidase family member and uses thereof
Patent: WO 0179473-A 1 25-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. 3143
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
229. .2877
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22990.1"
/db_xref="GI:18619527"

FEATURES
source
CDS

Query Match 99.5%; Score 3105.2; DB 6; Length 3143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 GCCTCGAGGCAAGCGCGCTGCTACTGCCCCCGCTGCTTCTAGTCCCGCTCCCGC 70
Db 26 GCCTCGAGGCAAGCGCGCTGCTACTGCCCCCGCTGCTTCTAGTCCCGCTCCCGC 85

QY 71 CTGGGTTGTCAACCGCGCGCGCGAGGAAAGCCACTGCAACAGGACCGAGTGGAGGC 130
Db 86 CTGGGTTGTCAACCGCGCGCGCGAGGAAAGCCACTGCAACAGGACCGAGTGGAGGC 145

QY 131 GGCGCAGCATGAAGCGCGCGAGCGCGCTCCATAGCGCACGTCGGGACGTCGCGCGCGG 190
Db 146 GGCGCAGCATGAAGCGCGCGAGCGCGCTCCATAGCGCACGTCGGGACGTCGCGCGCGG 205

QY 191 GCCCGGGGAAAGAAAATGCAACATGGCAGCAGCAATGGAACAGAAACAGTGGGTGTG 250
Db 206 GCCCGGGGAAAGAAAATGCAACATGGCAGCAGCAATGGAACAGAAACAGTGGGTGTG 265

QY 251 AGATATTTCAAACTCGGACTGTGAGGAGAAATTTGAATCAACAGATCGGCTAAATTTG 310
Db 266 AGATATTTCAAACTCGGACTGTGAGGAGAAATTTGAATCAACAGATCGGCTAAATTTG 325

QY 311 AGCCTTTTATGTTGAGCGGTATTCCTGAGTTCAGCTTAAAAAGCTGCTGCCGATACCA 370
Db 326 AGCCTTTTATGTTGAGCGGTATTCCTGAGTTCAGCTTAAAAAGCTGCTGCCGATACCA 385

ORIGIN

QY 371 GAAATATCATGGCTACATGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGA 430
DB |||||
QY 386 GAAATATCATGGCTACATGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGA 445
DB |||||
QY 431 ATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTACCTTGGCCATGCTGTGTGAACA 490
DB |||||
QY 446 ATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTACCTTGGCCATGCTGTGTGAACA 505
DB |||||
QY 491 GAGAAATACACTGTTTATTTCTGAAATTTCCAAATCTCAATAGACGACGAGTCTTAA 550
DB |||||
QY 506 GAGAAATACACTGTTTATTTCTGAAATTTCCAAATCTCAATAGACGACGAGTCTTAA 565
QY 551 TGCTCTTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATT 610
DB |||||
QY 566 TGCTCTTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATT 625
QY 611 CTCGAGAGAGAACTATTAAAGAGAGAGAAACCAATTTGGAACAGTCCGGAATTTGCTTTT 670
DB |||||
QY 626 CTCGAGAGAGAACTATTAAAGAGAGAGAAACCAATTTGGAACAGTCCGGAATTTGCTTTT 685
QY 671 ACATTTATCACCAAGAGAGTGAACATTTCTGTTTCAAGCGGCTAGTGGAAATTTATCACG 730
DB |||||
QY 686 ACATTTATCACCAAGAGAGTGAACATTTCTGTTTCAAGCGGCTAGTGGAAATTTATCACG 745
QY 731 TAAAGATGGAGGCGCCACAAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAA 790
DB |||||
QY 746 TAAAGATGGAGGCGCCACAAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAA 805
QY 791 CTAGTTGTCTCCAAACATACGATGGATTTCCAAATTTATGCCCCGCTGATCCAGACTGGATTG 850
DB |||||
QY 806 CTAGTTGTCTCCAAACATACGATGGATTTCCAAATTTATGCCCCGCTGATCCAGACTGGATTG 865
QY 851 CTTTATACATAGCAACGATTTTCGATATCTAACATCTGTAAACGAGAGAGAGAGAC 910
DB |||||
QY 866 CTTTATACATAGCAACGATTTTCGATATCTAACATCTGTAAACGAGAGAGAGAGAC 925
QY 911 TCATTTATGTCACATGAGTACGCAACATGGAAGAGATGCCAGATCAGCTGGAGTGG 970
DB |||||
QY 926 TCATTTATGTCACATGAGTACGCAACATGGAAGAGATGCCAGATCAGCTGGAGTGG 985
QY 971 CTACCTTTGTCTCCAAAGAGAAATTTGATAGATATCTGCGTATTTGGTGTCTCCAAAG 1030
DB |||||
QY 986 CTACCTTTGTCTCCAAAGAGAAATTTGATAGATATCTGCGTATTTGGTGTCTCCAAAG 1045
QY 1031 CTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAAATGATCAAT 1090
DB |||||
QY 1046 CTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAAATGATCAAT 1105
QY 1091 CTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGGAAACAAAGAGGCGAGATTCAT 1150
DB |||||
QY 1106 CTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGGAAACAAAGAGGCGAGATTCAT 1165
QY 1151 TCCGTTATCTCTAAACAGGTAAGCAATCTCTAAAGTCACTTTTAAAGATGTACAGAAATA 1210
DB |||||
QY 1166 TCCGTTATCTCTAAACAGGTAAGCAATCTCTAAAGTCACTTTTAAAGATGTACAGAAATA 1225
QY 1211 TGATTCATGCTGAAGAGGATCATAGATGTATAGATAAGAACTAATTTCAACCTTTTG 1270
DB |||||
QY 1226 TGATTCATGCTGAAGAGGATCATAGATGTATAGATAAGAACTAATTTCAACCTTTTG 1285
QY 1271 AGATTTCTTTTGAAGAGTTGAATATTTTGGCAGAGCTGGATGGACTCTCTGAGGAAAT 1330
DB |||||
QY 1286 AGATTTCTTTTGAAGAGTTGAATATTTTGGCAGAGCTGGATGGACTCTCTGAGGAAAT 1345
QY 1331 ATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGTATCTCAC 1390
DB |||||
QY 1346 ATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGTATCTCAC 1405
QY 1391 CTGAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGACTCATTTGAGTCA 1450
DB |||||
QY 1406 CTGAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGACTCATTTGAGTCA 1465
QY 1451 TGCTGTATTTCTGTAGCGCCACTAAATTTATCTATGAAGAGAAACACAGACATCTGGATAAATA 1510

DB |||||
QY 1466 TGCTGTATTTCTGTAGCGCCACTAAATTTATCTATGAAGAGAAACACAGACATCTGGATAAATA 1525
QY 1511 TCCATCAGACATCTTTTCATGTTTCCCAAAGTTCAGAGAGAGAAATTTGAGTTTATTTTTT 1570
DB |||||
QY 1526 TCCATCAGACATCTTTTCATGTTTCCCAAAGTTCAGAGAGAGAAATTTGAGTTTATTTTTT 1585
QY 1571 CCTCTGAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATTTTAAAGGAAA 1630
DB |||||
QY 1586 CCTCTGAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATTTTAAAGGAAA 1645
QY 1631 GCAAAATATAAACGATCCAGTGGTGGCTGCTTCCAAAGTGTTCAAAGTGTCTTATCA 1690
DB |||||
QY 1646 GCAAAATATAAACGATCCAGTGGTGGCTGCTTCCAAAGTGTTCAAAGTGTCTTATCA 1705
QY 1691 AAGAGAGATAGCAATTTACAGTGTGGAATGGGAAGTTCCTTGGCGGCGATGGATCTAATA 1750
DB |||||
QY 1706 AAGAGAGATAGCAATTTACAGTGTGGAATGGGAAGTTCCTTGGCGGCGATGGATCTAATA 1765
QY 1751 TCCAAAGTGTGATGAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAG 1810
DB |||||
QY 1766 TCCAAAGTGTGATGAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAG 1825
QY 1811 AGCATCACCTGTACGTAGTACGTACGTAAATTCCTGGAGAGGTGACAAAGGCTGACTGACC 1870
DB |||||
QY 1826 AGCATCACCTGTACGTAGTACGTAAATTCCTGGAGAGGTGACAAAGGCTGACTGACC 1885
QY 1871 GTGGCTACTCACATCTCTGCTCATCAGTACGACTGTGACTTCTTTATAGTAAGTATA 1930
DB |||||
QY 1886 GTGGCTACTCACATCTCTGCTCATCAGTACGACTGTGACTTCTTTATAGTAAGTATA 1945
QY 1931 GTAAACAGAGAAATCCACACTGTGTCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACC 1990
DB |||||
QY 1946 GTAAACAGAGAAATCCACACTGTGTCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACC 2005
QY 1991 CAACTTGCAGAAATTTTGGGCGCACATTTTGGGCGCACATTTTGGATTCAGCAGGCTCTCTCTG 2050
DB |||||
QY 2006 CAACTTGCAGAAATTTTGGGCGCACATTTTGGGCGCACATTTTGGATTCAGCAGGCTCTCTCTG 2065
QY 2051 ACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCTGGATTTACATTTGATGGGA 2110
DB |||||
QY 2066 ACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCTGGATTTACATTTGATGGGA 2125
QY 2111 TGCTCTACAGGCTCATGATCTACAGCTCGGAAGAAATATCTCTACTGCTGCTTTTCATAT 2170
DB |||||
QY 2126 TGCTCTACAGGCTCATGATCTACAGCTCGGAAGAAATATCTCTACTGCTGCTTTTCATAT 2185
QY 2171 ATGCTGTCTCTCAGGTGCAAGTGTGGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230
DB |||||
QY 2186 ATGCTGTCTCTCAGGTGCAAGTGTGGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2245
QY 2231 TGAATACCTTAGGCTCTCTAGGTATGTTGTAGTATAGACAAACAGGGGATCTCTGTC 2290
DB |||||
QY 2246 TGAATACCTTAGGCTCTCTAGGTATGTTGTAGTATAGACAAACAGGGGATCTCTGTC 2305
QY 2291 ACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAATAATGGGTCAAAATAGAAATGACG 2350
DB |||||
QY 2306 ACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAATAATGGGTCAAAATAGAAATGACG 2365
QY 2351 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTGACTTAGATCGTG 2410
DB |||||
QY 2366 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTGACTTAGATCGTG 2425
QY 2411 TGGGATCCACGGCTGCTATAGGAGTACTCTCTCCCTGATGTCATTTATGACAGGT 2470
DB |||||
QY 2426 TGGGATCCACGGCTGCTATAGGAGTACTCTCTCCCTGATGTCATTTATGACAGGT 2485
QY 2471 CAGATATCTTCAGGTTGCTATTTGCTGGGCGGCGGCTCACTCTGTTGGATCTTCTATGATA 2530
DB |||||
QY 2486 CAGATATCTTCAGGTTGCTATTTGCTGGGCGGCGGCTCACTCTGTTGGATCTTCTATGATA 2545
QY 2531 CAGGATACGGAAGCTTATATGGGTTCACCTTGACAGAAATGAACAGGCTTACTTACTTAG 2590

1153 CGTTATCTTAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTGAGAAATATG 1212
Db
1142 CGTTATCTTAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTGAGAAATATG 1201
Qy
1213 ATTGATGCTGAAGGAAGATCATAGATGTCTATAGATAAGGAACTTAATTCAAACCTTTTGAG 1272
Db
1202 ATTGATGCTGAAGGAAGATCATAGATGTCTATAGATAAGGAACTTAATTCAAACCTTTTGAG 1261
Qy
1273 ATTCTATTGGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAAT 1332
Db
1262 ATTCTATTGGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAAT 1321
Qy
1333 GCTTGGTGCATCTCTACTAGATCGCTCCAGAGCTGCCCTACAGATAGTGTGTCTCACCT 1392
Db
1322 GCTTGGTGCATCTCTACTAGATCGCTCCAGAGCTGCCCTACAGATAGTGTGTCTCACCT 1381
Qy
1393 GAATATTATTATCCAGTAGAAGATGATGTTATGAAAGGAGAGACTCAATGAGTCAGTG 1452
Db
1382 GAATATTATTATCCAGTAGAAGATGATGTTATGAAAGGAGAGACTCAATGAGTCAGTG 1441
Qy
1453 CTTGATTTCTGACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATC 1512
Db
1442 CTTGATTTCTGACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATC 1501
Qy
1513 CATGACATCTTTTCATGTTTTTCCCAAGTCCAGAGGAAATTCAGTGTATTTTGGC 1572
Db
1502 CATGACATCTTTTCATGTTTTTCCCAAGTCCAGAGGAAATTCAGTGTATTTTGGC 1561
Qy
1573 TCTGAATGCAAAAGGTTTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAGC 1632
Db
1562 TCTGAATGCAAAAGGTTTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAGC 1621
Qy
1633 AAATATAAACGATCAGTGGTGGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAA 1692
Db
1622 AAATATAAACGATCAGTGGTGGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAA 1681
Qy
1693 GAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGCGCGGATGATCTAATATC 1752
Db
1682 GAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGCGCGGATGATCTAATATC 1741
Qy
1753 CAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCAAAGACTCCCTTTAGAG 1812
Db
1742 CAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCAAAGACTCCCTTTAGAG 1801
Qy
1813 CATCACTGTAGTACAGTACAGTAACTCTGAGAGGTGCAAGGCTGACGTGACCGT 1872
Db
1802 CATCACTGTAGTACAGTACAGTAACTCTGAGAGGTGCAAGGCTGACGTGACCGT 1861
Qy
1873 GGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTTCTTTATAAGTAAAGTATAGT 1932
Db
1862 GGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTTCTTTATAAGTAAAGTATAGT 1921
Qy
1933 AACGAGAAATCCACACTGTGTCTTTCCTTTTAAAGCTATCAAGTCTGAGATGACCCA 1992
Db
1922 AACGAGAAATCCACACTGTGTCTTTCCTTTTAAAGCTATCAAGTCTGAGATGACCCA 1981
Qy
1993 ACTTGCAGAAACAAAGGAATTTTGGCCACCAATTTTGGATTCAGCAGCTCTCTTCTGAC 2052
Db
1982 ACTTGCAGAAACAAAGGAATTTTGGCCACCAATTTTGGATTCAGCAGCTCTCTTCTGAC 2041
Qy
2053 TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTCAATTTGATGGGATG 2112
Db
2042 TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTCAATTTGATGGGATG 2101
Qy
2113 CTCTACAGCCTCATGATCTACAGCCTGGGAAAGAAATATCTCTACTGTGCTGTTCAATAT 2172
Db
2102 CTCTACAGCCTCATGATCTACAGCCTGGGAAAGAAATATCTCTACTGTGCTGTTCAATAT 2161
Qy
2173 GGTGGTCTCAGTCCAGTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTG 2232
Db
2162 GGTGGTCTCAGTCCAGTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTG 2221
Qy
2233 AATACCTTAGCCTCTCTAGGTTATGTTGTTGATGTATAGACAAACAGGGGATCTCTGCAC 2292

2222 AATACCTTAGCCTCTCTAGGTTATGTTGTTGATGTATAGACAAACAGGGGATCTCTGCAC 2281
Qy
2293 CGAGGCTTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAATGACGAT 2352
Db
2282 CGAGGCTTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAATGACGAT 2341
Qy
2353 CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTTGATCTAGATCGGTG 2412
Db
2342 CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTTGATCTAGATCGGTG 2401
Qy
2413 GGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGTCAGAGGTCA 2472
Db
2402 GGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGTCAGAGGTCA 2461
Qy
2473 GATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACA 2532
Db
2462 GATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACA 2521
Qy
2533 GGATACACGGAGCTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGA 2592
Db
2522 GGATACACGGAGCTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGA 2581
Qy
2593 TCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACAT 2652
Db
2582 TCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACAT 2641
Qy
2653 GGTTCCTCGATGAGAAATGTCATTTTGCACATACAGTATATTTACTGAGTTTTTTAGTG 2712
Db
2642 GGTTCCTCGATGAGAAATGTCATTTTGCACATACAGTATATTTACTGAGTTTTTTAGTG 2701
Qy
2713 AGGGCTGGAAGCAGATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGATT 2772
Db
2702 AGGGCTGGAAGCAGATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGATT 2761
Qy
2773 CCTGAATCGGAGAACATTTATGAATGCTATCTTTTGGACTACTCTTCAAGAAACCTTTGGA 2832
Db
2762 CCTGAATCGGAGAACATTTATGAATGCTATCTTTTGGACTACTCTTCAAGAAACCTTTGGA 2821
Qy
2833 TCACGTATTGCTGCTCTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTCGGTATAC 2892
Db
2822 TCACGTATTGCTGCTCTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTCGGTATAC 2881
Qy
2893 ACTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCAC 2952
Db
2882 ACTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCAC 2941
Qy
2953 ATTTTGATACCTGCGCATGTAAACATCTACTCTGAAATAAATGTGGTCCCATGAGGGGT 3012
Db
2942 ATTTTGATACCTGCGCATGTAAACATCTACTCTGAAATAAATGTGGTCCCATGAGGGGT 3001
Qy
3013 CTACGGTTTGTGTAGTAAATCTAAATCACTTAAACCCACATGCTCAAAATCAAAATGATACA 3072
Db
3002 CTACGGTTTGTGTAGTAAATCTAAATCACTTAAACCCACATGCTCAAAATCAAAATGATACA 3061
Qy
3073 TATTCTGAGAGAGCCAGCAATACCATAAGAAATTAATAAAAAA 3117
Db
3062 TATTCTGAGAGAGCCAGCAATACCATAAGAAATTAATAAAAAA 3106

RESULT 7
BC040203

LOCUS
DEFINITION Homo sapiens dipeptidylpeptidase 8, mRNA (cdna clone MGC:42570

IMAGE:4824813), complete cds.

ACCESSION BC040203

VERSION BC040203.1 GI:26007916

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE AUTHORS	1 (bases 1 to 4535) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Cabavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y., Skrzewinski,M.I., Skalska,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences (26), 16899-16903 (2002) Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932
JOURNAL PUBMED	2 (bases 1 to 4535) Strausberg,R. Direct Submission
REFERENCE AUTHORS TITLE	Submitted (27-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleaged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Wuzny, D.M., Nanavati, A.N., Gibbs, R.A.
REMARK COMMENT	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 71 Row: K Column: 23. Location/Qualifiers 1. .4535 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:42570 IMAGE:4824813" /tissue_type="Testis" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Vector: pBluescript" 1. .4535 /gene="dpp8" /note="synonyms: FLJ20283, FLJ14920, MGC26191" /db_xref="GeneID:54878" /db_xref="MIM:606819" 1629. .4277 /gene="dpp8" /codon_start=1 /product="dpp8 protein" /protein_id="AAH40203.1" /db_xref="GI:26007917" /db_xref="GeneID:54878" /db_xref="MIM:606819" /translation="MAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQ
FEATURES source	1. .4535 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:42570 IMAGE:4824813" /tissue_type="Testis" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Vector: pBluescript" 1. .4535 /gene="dpp8" /note="synonyms: FLJ20283, FLJ14920, MGC26191" /db_xref="GeneID:54878" /db_xref="MIM:606819" 1629. .4277 /gene="dpp8" /codon_start=1 /product="dpp8 protein" /protein_id="AAH40203.1" /db_xref="GI:26007917" /db_xref="GeneID:54878" /db_xref="MIM:606819" /translation="MAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQ
gene	1. .4535 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:42570 IMAGE:4824813" /tissue_type="Testis" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Vector: pBluescript" 1. .4535 /gene="dpp8" /note="synonyms: FLJ20283, FLJ14920, MGC26191" /db_xref="GeneID:54878" /db_xref="MIM:606819" 1629. .4277 /gene="dpp8" /codon_start=1 /product="dpp8 protein" /protein_id="AAH40203.1" /db_xref="GI:26007917" /db_xref="GeneID:54878" /db_xref="MIM:606819" /translation="MAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQ
CDS	1. .4535 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:42570 IMAGE:4824813" /tissue_type="Testis" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Vector: pBluescript" 1. .4535 /gene="dpp8" /note="synonyms: FLJ20283, FLJ14920, MGC26191" /db_xref="GeneID:54878" /db_xref="MIM:606819" 1629. .4277 /gene="dpp8" /codon_start=1 /product="dpp8 protein" /protein_id="AAH40203.1" /db_xref="GI:26007917" /db_xref="GeneID:54878" /db_xref="MIM:606819" /translation="MAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQ
misc_feature	1. .4535 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:42570 IMAGE:4824813" /tissue_type="Testis" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Vector: pBluescript" 1. .4535 /gene="dpp8" /note="synonyms: FLJ20283, FLJ14920, MGC26191" /db_xref="GeneID:54878" /db_xref="MIM:606819" 1629. .4277 /gene="dpp8" /codon_start=1 /product="dpp8 protein" /protein_id="AAH40203.1" /db_xref="GI:26007917" /db_xref="GeneID:54878" /db_xref="MIM:606819" /translation="MAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQ
ORIGIN	Query Match 93.4%; Score 2915.2; DB 8; Length 4535; Best Local Similarity 99.9%; Pred. No. 0; Matches 2917; Conservative 0; Mismatches 3; Indels 0; Gaps 0; Qy 201 AGMAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGTGTTCAGATATTTGA 260 Db 1616 AGMAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGTGTTCAGATATTTGA 1675 Qy 261 AACTGGGACTGTGAGGAGAAATGATACAGATCGGCCTAAATGGAGCCTTTTGA 320 Db 1676 AACTGGGACTGTGAGGAGAAATGATACAGATCGGCCTAAATGGAGCCTTTTGA 1735 Qy 321 TGTGAGCGTATTCCTGGAGTCAGCTTAAAGCTGCTCCGATACAGAAATATCA 380 Db 1736 TGTGAGCGTATTCCTGGAGTCAGCTTAAAGCTGCTCCGATACAGAAATATCA 1795 Qy 381 TGCTACATGATGGCTAAGGCCACCATGATTTTCAATGTTGTGAAGAGGAATGATCCAGA 440 Db 1796 TGCTACATGATGGCTAAGGCCACCATGATTTTCAATGTTGTGAAGAGGAATGATCCAGA 1855 Qy 441 TGGACCTCATTGAGACAGAGATCTTATACCTTGGCATGTCTGGTGAAGACAGAGAAATAC 500 Db 1856 TGGACCTCATTGAGACAGAGATCTTATACCTTGGCATGTCTGGTGAAGACAGAGAAATAC 1915 Qy 501 ACTGTTTATTCGTAATTCCTGAAATTCCTGAAATTCCTGAAATTCCTGAAATTCCTG 560 Db 1916 ACTGTTTATTCGTAATTCCTGAAATTCCTGAAATTCCTGAAATTCCTGAAATTCCTG 1975 Qy 561 GAAGCCTCTTTTGGATCTTTTTCAGGCAACATCGACTATGGAATGATTTCTCGAGAAGA 620 Db 1976 GAAGCCTCTTTTGGATCTTTTTCAGGCAACATCGACTATGGAATGATTTCTCGAGAAGA 2035 Qy 621 AGAAGCTTTAAG 680 Db 2036 AGAAGCTTTAAG 2095 Qy 681 CCAAG 740 Db 2096 CCAAG 2155 Qy 741 AGGGCCCAAG 800 Db 2156 AGGGCCCAAG 2215 Qy 801 CAACATACCGAG 860 Db 2216 CAACATACCGAG 2275 Qy 861 TAGCAACGATATTTTGGATATCTAACATCGTAAACAGAGAGAGAGAGAGAGAGAGAG 920

Db 2276 TGGCAACGATATTTGGATATCTAAATCGTAAACAGAGAAAGGAGACTCACTATGT 2335
Qy 921 GCACATAGCTAGCCAAATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 980
Db 2336 GCACATAGCTAGCCAAATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 2395
Qy 981 TCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTATTTGGTGTGTCGCTACCTTTGT 1040
Db 2396 TCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTATTTGGTGTGTCGCTACCTTTGT 2455
Qy 1041 TCCCAAGTGTGTAAGAAATCTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGA 1100
Db 2456 TCCCAAGTGTGTAAGAAATCTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGA 2515
Qy 1101 AATTATTCATGTTACATCCCTATGTTTGGAAACAAGGAGGCGAGATTCATTCGGTATCC 1160
Db 2516 AATTATTCATGTTACATCCCTATGTTTGGAAACAAGGAGGCGAGATTCATTCGGTATCC 2575
Qy 1161 TAAAAACAGGTACAGCAAAATCCTAAAGTCATCTTTAAGATGTCAGAAATATGATGATGC 1220
Db 2576 TAAAAACAGGTACAGCAAAATCCTAAAGTCATCTTTAAGATGTCAGAAATATGATGATGC 2635
Qy 1221 TGAAGGAGGATCATAGATGTCATAGATTAAGGAACTAAATCAACCTTTTGAGATTCAT 1280
Db 2636 TGAAGGAGGATCATAGATGTCATAGATTAAGGAACTAAATCAACCTTTTGAGATTCAT 2695
Qy 1281 TGAAGGAGGATTAATATTTGCAGAGCTGGATGACTCCTGAGGAAATATGCTTGGTC 1340
Db 2696 TGAAGGAGGATTAATATTTGCAGAGCTGGATGACTCCTGAGGAAATATGCTTGGTC 2755
Qy 1341 CATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAACCTGAATAT 1400
Db 2756 CATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAACCTGAATAT 2815
Qy 1401 TATCCAGTAGAAGATGATTTTGAAGAGCAGAGACTCAATGAGTCAGTCACTGATTC 1460
Db 2816 TATCCAGTAGAAGATGATTTTGAAGAGCAGAGACTCAATGAGTCAGTCACTGATTC 2875
Qy 1461 TGTGAGCAGCTAAATTAATCTATGAAGAAACAAGACATCTGGATAAATATCATGACAT 1520
Db 2876 TGTGAGCAGCTAAATTAATCTATGAAGAAACAAGACATCTGGATAAATATCATGACAT 2935
Qy 1521 CTTTCATGTTTTCCTCCAAAGTCAGAGAGGAAATGAGTTTATTTTTCCTCTGAATG 1580
Db 2936 CTTTCATGTTTTCCTCCAAAGTCAGAGAGGAAATGAGTTTATTTTTCCTCTGAATG 2995
Qy 1581 CAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAA 1640
Db 2996 CAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAA 3055
Qy 1641 ACGATCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAGTGTCTATCAAGAGGAGAT 1700
Db 3056 ACGATCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAGTGTCTATCAAGAGGAGAT 3115
Qy 1701 AGCAATACAGTGGTGAATGGGAATTTCTGGCCGCGATGGATCTAAATATCAAGTTGA 1760
Db 3116 AGCAATACAGTGGTGAATGGGAATTTCTGGCCGCGATGGATCTAAATATCAAGTTGA 3175
Qy 1761 TGAAGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGATCCTCCCTTTAGAGATCACCT 1820
Db 3176 TGAAGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGATCCTCCCTTTAGAGATCACCT 3235
Qy 1821 GTACGTAGTCAGTTAGTAAATCTTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTC 1880
Db 3236 GTACGTAGTCAGTTAGTAAATCTTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTC 3295
Qy 1881 ACATTCCTGCTGCATCAGTCAGTCACTGTGACTCTCTTATAAGTAAGTATAGTAACAGAA 1940
Db 3296 ACATTCCTGCTGCATCAGTCAGTCACTGTGACTCTCTTATAAGTAAGTATAGTAACAGAA 3355
Qy 1941 GAATCCACACTGTGTGCTCCTTTTAAAGGATCAAGTCTCAAGTCTGAGAGTGAACCACTTGC 2000
Db 3356 GAATCCACACTGTGTGCTCCTTTTAAAGGATCAAGTCTCAAGTCTGAGAGTGAACCACTTGC 3415

Qy 2001 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTTCTGACTATCTCC 2060
Db 3416 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTTCTGACTATCTCC 3475
Qy 2061 TCCAGAAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAA 2120
Db 3476 TCCAGAAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAA 3535
Qy 2121 GCCTCATGATCTACAGCCTGGAAATAATCTCTACTGTCTGTTTCAATATATGTTGTTCC 2180
Db 3536 GCCTCATGATCTACAGCCTGGAAATAATCTCTACTGTCTGTTTCAATATATGTTGTTCC 3595
Qy 2181 TCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 2240
Db 3596 TCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 3655
Qy 2241 AGCCTCTCTAGGTTATGTGTTGTAGTATGATACAAACAGGGGATCTGTCTCACCCGAGGCT 2300
Db 3656 AGCCTCTCTAGGTTATGTGTTGTAGTATGATACAAACAGGGGATCTGTCTCACCCGAGGCT 3715
Qy 2301 TAAATTTGAAGCGCCTTTTAAATATATAAATGGGTCAAAATAGAAAATTCACATCAGGTGGA 2360
Db 3716 TAAATTTGAAGCGCCTTTTAAATATATAAATGGGTCAAAATAGAAAATTCACATCAGGTGGA 3775
Qy 2361 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGTGGGCATCCA 2420
Db 3776 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGTGGGCATCCA 3835
Qy 2421 CGGCTGGTCTATGAGGATACCTCTCCCTGATGTCATTAATGAGGATTCAGAGGTGATATCTT 2480
Db 3836 CGGCTGGTCTATGAGGATACCTCTCCCTGATGTCATTAATGAGGATTCAGAGGTGATATCTT 3895
Qy 2481 CAGGTTGCTATTTGCTGGGGCCCGCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 2540
Db 3896 CAGGTTGCTATTTGCTGGGGCCCGCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 3955
Qy 2541 GGAACGTTATATGGGTACCTTGACAGAAATGAACAGGGCTATTTACTTGAATCTGTGGC 2600
Db 3956 GGAACGTTATATGGGTACCTTGACAGAAATGAACAGGGCTATTTACTTGAATCTGTGGC 4015
Qy 2601 CATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGCTTTCCT 2660
Db 4016 CATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGCTTTCCT 4075
Qy 2661 GGATGAGAAATGTCATTTTGGACATACACAGTATATTTACTGAGTTTCTTGTAGGGCTGG 2720
Db 4076 GGATGAGAAATGTCATTTTGGACATACACAGTATATTTACTGAGTTTCTTGTAGGGCTGG 4135
Qy 2721 AAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTGAAATC 2780
Db 4136 AAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTGAAATC 4195
Qy 2781 GGGAGAACATTTATGAACTGCTGATCTTTTGGATCTTCAAGAAACCTTGGATCAGGTAT 2840
Db 4196 GGGAGAACATTTATGAACTGCTGATCTTTTGGATCTTCAAGAAACCTTGGATCAGGTAT 4255
Qy 2841 TGCTGCTCTAAAGTATATTTTGGCTGTGAGAACTCTCTGTTGATACATCTGGCTA 2900
Db 4256 TGCTGCTCTAAAGTATATTTTGGCTGTGAGAACTCTCTGTTGATACATCTGGCTA 4315
Qy 2901 TTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTAT 2960
Db 4316 TTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTAT 4375
Qy 2961 ACCTGCCATGTAAATCTACTCTGAAATAAATGTTGGTGCATGACAGGGGTCTACGGTT 3020
Db 4376 ACCTGCCATGTAAATCTACTCTGAAATAAATGTTGGTGCATGACAGGGGTCTACGGTT 4435
Qy 3021 TGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATTTCTG 3080
Db 4436 TGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATTTCTG 4495

Qy	3081	AGAGACCCAGCAATACCATAGAAATTAATAAAAAAAAAA	3120																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
----	------	---	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

QY	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAAGCTGGTATATATTTTGAAGGCCAACAGAC	1800
Db	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAAGCTGGTATATATTTTGAAGGCCAACAGAC	1800
QY	1801	TCCCTTTAGAGCATACCTGTAGCTAGTACGTTAGCTAAATCCTGGAGAGGTACAAGG	1860
Db	1801	TCCCTTTTAGAGCATACCTGTAGCTAGTACGTTAGCTAAATCCTGGAGAGGTACAAGG	1860
QY	1861	CTGACTGACCGTGTACTACATCTCTGCTGCATCAGTCAGCACTGTGACTCTTTTATA	1920
Db	1861	CTGACTGACCGTGTACTACATCTCTGCTGCATCAGTCAGCACTGTGACTCTTTTATA	1920
QY	1921	AGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCT	1980
Db	1921	AGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCT	1980
QY	1981	GAAATGACCCAACTTGCAAAACAAAGGAAATTTGGGCCACCATTTGGATTACAGAGT	2040
Db	1981	GAAATGACCCAACTTGCAAAACAAAGGAAATTTGGGCCACCATTTGGATTACAA-----	2034
QY	2041	CCTCTTCTGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA	2100
Db	2035	-----	2034
QY	2101	TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTCGAAGAAATATCTTACTGTG	2160
Db	2035	-----	2034
QY	2161	CTGTTCAATATATGTTGTCCTCAGGTGCGAGTTGGTGATATATCGGTTTAAAGGAGTCAAG	2220
Db	2035	-----GTCCCTCAGGTGCGAGTTGGTGATATATCGGTTTAAAGGAGTCAAG	2078
QY	2221	TATTTCCGCTTGATACCTCTAGCTCTAGTTAGTTAGTTAGTATAGACACACAGG	2280
Db	2079	TATTTCCGCTTGATACCTCTAGCTCTCTAGGTTATGTTGTTAGTATAGACACACAGG	2138
QY	2281	GGATCCTGTCTACCGAGGCTTAAATTTGAAGGCGCTTAAATATATAAAATGGGTCAATA	2340
Db	2139	GGATCCTGTCTACCGAGGCTTAAATTTGAAGGCGCTTAAATATATAAAATGGGTCAATA	2198
QY	2341	GAATATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2400
Db	2199	GAATATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2258
QY	2401	TTAGATCTGTGGGCATCCAGCGTGTCTATGAGGATACCTCTCCCTGATGGCATT	2460
Db	2259	TTAGATCTGTGGGCATCCAGCGTGTCTATGAGGATACCTCTCCCTGATGGCATT	2318
QY	2461	ATGCAGAGGTCAGATATCTTCAGGTTGCTTATGTCGGGCGCCAGTCACTCTGTGGATC	2520
Db	2319	ATGCAGAGGTCAGATATCTTCAGGTTGCTTATGTCGGGCGCCAGTCACTCTGTGGATC	2378
QY	2521	TTCTATGATACAGGATACCGAAAGCTTATATGAGGTCACCTTGACAGAAATGAACAGGGC	2580
Db	2379	TTCTATGATACAGGATACCGAAAGCTTATATGAGGTCACCTTGACAGAAATGAACAGGGC	2438
QY	2581	TATTAATTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAACCAATCGTTTA	2640
Db	2439	TATTAATTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAACCAATCGTTTA	2498
QY	2641	CTGCTCTTACATGTTTCTGGATGAGATGTCATTTTGGCATACACAGTATATTACTG	2700
Db	2499	CTGCTCTTACATGTTTCTGGATGAGATGTCATTTTGGCATACACAGTATATTACTG	2558
QY	2701	AGTTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACAC	2760
Db	2559	AGTTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACAC	2618
QY	2761	AGCATPAGAGTTCTGAAATCGGGAGAACATTATGAACTGTGATCTTTTGTGACATACCTTCA	2820
Db	2619	AGCATPAGAGTTCTGAAATCGGGAGAACATTATGAACTGTGATCTTTTGTGACATACCTTCA	2678
QY	2821	GAATACTTGGATCAGTATCTGCTCTAAAGTGATATATAATTTTGACCTGTGTAGAAC	2880
Db	2679	GAATACTTGGATCAGTATCTGCTCTAAAGTGATATATAATTTTGACCTGTGTAGAAC	2738
QY	2881	TCTCTGTATACACTCGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAG	2940
Db	2739	TCTCTGTATACACTCGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAG	2798
QY	2941	ATTGATCATCACATTTTGTATACCTGCGATGTAACATCTACTCTCGAATAAATATGTGGTG	3000
Db	2799	ATTGATCATCACATTTTGTATACCTGCGATGTAACATCTACTCTCGAATAAATATGTGGTG	2858
QY	3001	CCATGAGGCGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAA	3060
Db	2859	CCATGAGGCGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAA	2918
QY	3061	TCAAATCATACATATCTCTGAGAGACCCAGCAATACCAATAAGAAATTAATAAAAAAAA	3120
Db	2919	TCAAATCATACATATCTCTGAGAGACCCAGCAATACCAATAAGAAATTAATAAAAAAAA	2978
RESULT 9			
AX608745			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2977; Conservative			
0; Mismatches			
1; Indels			
142; Gaps			
QY	1	AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTAGTGCGG	60
QY	61	CGTTGCGCGCTGGTGTGTACCGCGCGCGCGCGGAGGAGCCACTCAACCCAGGACCG	120
Db	61	CGTTGCGCGCTGGTGTGTACCGCGCGCGCGCGGAGGAGCCACTCAACCCAGGACCG	120
QY	121	GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGACGCTCGGGACGG	180
Db	121	GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGACGCTCGGGACGG	180
QY	181	TCCGCGCGCGCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAAACAG	240
Db	181	TCCGCGCGCGCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAAACAG	240
QY	241	CTGGGTGTGTAGATATTTGAAACTGCGGACTGTGAGGAGCAATATTGAATCAGAGATCGG	300
Db	241	CTGGGTGTGTAGATATTTGAAACTGCGGACTGTGAGGAGCAATATTGAATCAGAGATCGG	300
QY	301	CCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTTCTGAGTCAGCTTAAAGGCTGCTT	360
Db	301	CCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTTCTGAGTCAGCTTAAAGGCTGCTT	360

Qy	361	GCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCACACATGATTTTCATGTTT	420	Db	1441	ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAAATTAATCTATGAAGAAACAACAGACATC	1500
Db	361	GCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCACACATGATTTTCATGTTT	420	Qy	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCACGAGAGGGAATTCAG	1560
Qy	421	GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCT	480	Db	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCACGAGAGGGAATTCAG	1560
Db	421	GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCT	480	Qy	1561	TTTATTTTTCCTCTGAATGCAAAAACAGGTTTCCTGTCATTTTATACAAAATACATCTATT	1620
Qy	481	GCTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA	540	Db	1561	TTTATTTTTCCTCTGAATGCAAAAACAGGTTTCCTGTCATTTTATACAAAATACATCTATT	1620
Db	481	GCTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA	540	Qy	1621	TTAAAGGAAACAAATATAAACGATCCAGTGGTGGCTGCCTCCCAAGTGAATTCGAG	1680
Qy	541	GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCACTGCACTAT	600	Db	1621	TTAAAGGAAACAAATATAAACGATCCAGTGGTGGCTGCCTCCCAAGTGAATTCGAG	1680
Db	541	GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCACTGCACTAT	600	Qy	1681	TGTCCTATCAAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Qy	601	GGAATGTATTCGAGAGAAAGAACTATTAAAGAGAAAGAAACGCATTTGGACAGTCGGA	660	Db	1681	TGTCCTATCAAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	601	GGAATGTATTCGAGAGAAAGAACTATTAAAGAGAAAGAAACGCATTTGGACAGTCGGA	660	Qy	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Qy	661	ATTGCTTCTTACGATTTATCACCAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA	720	Db	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Db	661	ATTGCTTCTTACGATTTATCACCAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA	720	Qy	1801	TCGCCCTTTAGAGCATCACCTGTAGTCAGTTACGTAAATCCTCGAGAGGTGACAAGG	1860
Qy	721	ATTATATCAGTAAAAAGATGGAGGCCCAAGGATTTTACGCAACAACTTTTAAAGGCCAAT	780	Db	1801	TCGCCCTTTAGAGCATCACCTGTAGTCAGTTACGTAAATCCTCGAGAGGTGACAAGG	1860
Db	721	ATTATATCAGTAAAAAGATGGAGGCCCAAGGATTTTACGCAACAACTTTTAAAGGCCAAT	780	Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTCATCAGTCAGCACGTGCTTCTTTATA	1920
Qy	781	CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAAAATTATGCCCGCTGATCCA	840	Db	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTCATCAGTCAGCACGTGCTTCTTTATA	1920
Db	781	CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAAAATTATGCCCGCTGATCCA	840	Qy	1921	AGTAAGTATAGTAAACCAGAAATCCACATGTCGTGTCATCAGTCAGCACGTGCTTCTTTATA	1980
Qy	841	GACTGGATTTCTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAA	900	Db	1921	AGTAAGTATAGTAAACCAGAAATCCACATGTCGTGTCATCAGTCAGCACGTGCTTCTTTATA	1980
Db	841	GACTGGATTTCTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAA	900	Qy	1981	GAAGATGACCCAACTTGGCAAAACAAAGGAATTTTGGCCACCACTTTTGGATTC	2040
Qy	901	GAAGAGACTCAGTATGTCACAAATGATGCTAGCCCAACATCGNAGAGATGCCATCA	960	Db	1981	GAAGATGACCCAACTTGGCAAAACAAAGGAATTTTGGCCACCACTTTTGGATTC	2040
Db	901	GAAGAGACTCAGTATGTCACAAATGATGCTAGCCCAACATCGNAGAGATGCCATCA	960	Qy	2041	CCTCTTCTCGACTATATCTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGATTTACA	2100
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAGAAATTTGATAGATATCTGCTATTGGTGG	1020	Db	2035	-----	2034
Db	961	GCTGGAGTCGCTACCTTTGTTCTCCAGAAATTTGATAGATATCTGCTATTGGTGG	1020	Qy	2101	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Qy	1021	TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTATATGAAGAA	1080	Db	2035	-----	2034
Db	1021	TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTATATGAAGAA	1080	Qy	2161	CTGTTTCATATATGTTGGTCTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGAGGTCAAG	2220
Qy	1081	AATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAAAACAGGAGG	1140	Db	2035	-----GTCCTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGAGGTCAAG	2078
Db	1081	AATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAAAACAGGAGG	1140	Qy	2221	TATTTTCOGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTGTTAGTAGACAAAGG	2280
Qy	1141	GCAGATTCATTCGGTTATCCTAAAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAGATG	1200	Db	2079	TATTTTCOGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTGTTAGTAGACAAAGG	2138
Db	1141	GCAGATTCATTCGGTTATCCTAAAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAGATG	1200	Qy	2281	GGATCTGTCAACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2340
Qy	1201	TCAGAAATATGATTCGATGCTGAAGAGGATCATAGATGTCTAGAGATAAGGAACTAAT	1260	Db	2139	GGATCTGTCAACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2198
Db	1201	TCAGAAATATGATTCGATGCTGAAGAGGATCATAGATGTCTAGAGATAAGGAACTAAT	1260	Qy	2341	GAATTTGAACGATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTCAATGAC	2400
Qy	1261	CAACCTTTTCAGATTTCTATTGAGGAGTTGAATATATGCGCAGAGCTGGATGGACTCCT	1320	Db	2199	GAATTTGAACGATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTCAATGAC	2258
Db	1261	CAACCTTTTCAGATTTCTATTGAGGAGTTGAATATATGCGCAGAGCTGGATGGACTCCT	1320	Qy	2401	TTAGATCGTGGGCAATCCAGCGCTGCTATGCGAGGATACCTCTCCCTCATGCAATTA	2460
Qy	1321	GAGGAAATATCTGTTGGTCCATCTCTACTAGATGCTGCCAGACTCGCTACAGATAGTG	1380	Db	2259	TTAGATCGTGGGCAATCCAGCGCTGCTATGCGAGGATACCTCTCCCTCATGCAATTA	2318
Db	1321	GAGGAAATATCTGTTGGTCCATCTCTACTAGATGCTGCCAGACTCGCTACAGATAGTG	1380	Qy	2461	ATGCAAGGTGAGATATCTTTCAGGGTGTCTATTTGCTGGGGCCCAAGTCACCTCTGTGATC	2520
Qy	1381	TTGATCTCAGCTGAATTTTATCCAGTAGAGATGATGTTATGGAAGGACAGAGACTC	1440	Db	2319	ATGCAAGGTGAGATATCTTTCAGGGTGTCTATTTGCTGGGGCCCAAGTCACCTCTGTGATC	2378
Db	1381	TTGATCTCAGCTGAATTTTATCCAGTAGAGATGATGTTATGGAAGGACAGAGACTC	1440	Qy	2521	TTCTATGATACAGGATACAGGAACTTATATGGGTACCTCTGACCAAGAAAGACAGGCG	2580
Qy	1441	ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAAATTAATCTATGAAGAAACAACAGACATC	1500				

```
Db 2379 TTCTATGATACAGGATACAGGAACGTTTATATGGGTACCCCTGACCAAGATGAACAGGCG 2438
Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Db 2439 TATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTA 2498
Qy 2641 CTGCTCTTACATGCTTCTCTGATGAGATGTCCATTTTGCACATACCAAGTATATTACTG 2700
Db 2499 CTGCTCTTACATGCTTCTCTGATGAGATGTCCATTTTGCACATACCAAGTATATTACTG 2558
Qy 2701 AGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACAC 2760
Db 2559 AGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACAC 2618
Qy 2761 AGCATAAGAGTTCTTGAATCGGAGAACATTAATGAATGCAATCTTTTGCACCTACCTCAA 2820
Db 2619 AGCATAGAGTTCTTGAATCGGAGAACATTAATGAATGCAATCTTTTGCACCTACCTCAA 2678
Qy 2821 GAAAACTTTGGATACGTAATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
Db 2679 GAAAACTTTGGATACGTAATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAAC 2738
Qy 2881 TCTCTGGTATACATGCTGCTATTAACCAATGAGAGGTTTAATCAACAGAAACACAGA 2940
Db 2739 TCTCTGGTATACATGCTGCTATTAACCAATGAGAGGTTTAATCAACAGAAACACAGA 2798
Qy 2941 ATTGATCATCAATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAATGTGGTG 3000
Db 2799 ATTGATCATCAATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAATGTGGTG 2858
Qy 3001 CCATGCGGGGCTACGGTTTGGTGTAGTAATCTTAACCTTAACCCACATGCTCAAAA 3060
Db 2859 CCATGCGGGGCTACGGTTTGGTGTAGTAATCTTAACCTTAACCCACATGCTCAAAA 2918
Qy 3061 TCAATGATACATATTCCTGAGAGCCAGCAATACCAATGAGTAATTAATAAAAAAAA 3120
Db 2919 TCAATGATACATATTCCTGAGAGCCAGCAATACCAATGAGTAATTAATAAAAAAAA 2978

RESULT 10
AR631287
LOCUS AR631287 4676 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 20 from patent US 6844180.
ACCESSION AR631287
VERSION AR631287.1 GI:59770934
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4676)
AUTHORS Oi,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 20 18-JAN-2005;
Ferring BV;;
NLX;
FEATURES
    source
        1..4676
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Query Match 89.9%; Score 2806.4; DB 6; Length 4676;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;
Qy 1 AAGTGCTAAAGGCTCCGAGGCCAAGCGCGTCTACTGCCCGCGCTCTTTCTTAGTGCCG 60
Db 1 AAGTGCTAAAGGCTCCGAGGCCAAGCGCGTCTACTGCCCGCGCTCTTTCTTAGTGCCG 60
Qy 61 CGTTCCGCCCTGGGTTGTACCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCG 120
Db 61 CGTTCCGCCCTGGGTTGTACCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCG 120
```

```
Qy 121 GAGTGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGACGG 180
Db 121 GAGTGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGACGG 180
Qy 181 TCCGGCGCGCGCGCGGAGAAATGCAACATGCGCAGCAGCAATGGAAACAGACAG 240
Db 181 TCCGGCGCGCGCGCGGAGAAATGCAACATGCGCAGCAGCAATGGAAACAGACAG 240
Qy 241 CTGGGTGTTGAGATATTGAAACTCGGACCTGTGAGGAGATATTGATCAACAGATCGG 300
Db 241 CTGGGTGTTGAGATATTGAAACTCGGACCTGTGAGGAGATATTGATCAACAGATCGG 300
Qy 301 CCTAAATTCGAGCTTTTATGTTGAGCGGTATTCTCGAGTCAGCTTTAAAGACGCTGT 360
Db 301 CCTAAATTCGAGCTTTTATGTTGAGCGGTATTCTCGAGTCAGCTTTAAAGACGCTGT 360
Qy 361 GCCGATACAGAAAAATATCATGGCTACATGATGGCTAAGGCCACCAATGATTTTCTGTT 420
Db 361 GCCGATACAGAAAAATATCATGGCTACATGATGGCTAAGGCCACCAATGATTTTCTGTT 420
Qy 421 GTGAAAGGAATGATCCAGATGGACCTCATTTGAGACAGATCTATTACCTTGGCATGCT 480
Db 421 GTGAAAGGAATGATCCAGATGGACCTCATTTGAGACAGATCTATTACCTTGGCATGCT 480
Qy 481 GGTGAGACAGAGAAAAATACCTGTTTATCTGAAATTCCTCAAACTATCAATAGACGA 540
Db 481 GGTGAGACAGAGAAAAATACCTGTTTATCTGAAATTCCTCAAACTATCAATAGACGA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTATTCTCGAAGAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTCGA 660
Db 601 GGAATGTATTCTCGAAGAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTCGA 660
Qy 661 ATTGCTTTCAATGATATCCAGAGAGTGAAACATTTCTGTTTCAAGCCGCTAGTGA 720
Db 661 ATTGCTTTCAATGATATCCAGAGAGTGAAACATTTCTGTTTCAAGCCGCTAGTGA 720
Qy 721 ATTTATCAGTAAAGATGAGGGCCACAAAGATTTTACGCAACACCTTTAAGGCCCAAT 780
Db 721 ATTTATCAGTAAAGATGAGGGCCACAAAGATTTTACGCAACACCTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAACCTAGTTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAACCTAGTTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA 840
Qy 841 GACTGGAATGCTTTTATACATAGCAACGATATTGGATATCTAACATCGTAACAGAGAA 900
Db 841 GACTGGAATGCTTTTATACATAGCAACGATATTGGATATCTAACATCGTAACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTCACATGAGCTAGCCACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTCACATGAGCTAGCCACATGGAAGAGATGCCAGATCA 960
Qy 961 GCTGGAGTCTGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCTGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTAATAATTTCTAGAAATCTATATAGAGAA 1080
Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTAATAATTTCTAGAAATCTATATAGAGAA 1080
Qy 1081 AATGATGAATCTGAGTGGAATTTATTCATGTTATACCTCCCTATGTTGGAACAGGAGG 1140
Db 1081 AATGATGAATCTGAGTGGAATTTATTCATGTTATACCTCCCTATGTTGGAACAGGAGG 1140
Qy 1141 GCAGATTCATTCCGTTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCATTCCGTTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200
Qy 1201 TCAGAAATATGATTTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACTAAT 1260
```


source	1. .4676																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
--------	----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Qy	2041	CCTCTTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
Db	2041	CCTCTTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
Qy	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAATATCTCTACTGTG	2160
Db	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAATATCTCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTTGTCCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG	2220
Db	2161	CTGTTTCATATATGTTGTCCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG	2179
Qy	2221	TATTTCCGCTTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGG	2280
Db	2180	-----	2179
Qy	2281	GGATCCTGTGACCAGGGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAATA	2340
Db	2180	-----GGGTCAATAA	2189
Qy	2341	GAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTCAC	2400
Db	2190	GAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTCAC	2249
Qy	2401	TTAGATCTGTGGGCATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGCAATTA	2460
Db	2250	TTAGATCTGTGGGCATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGCAATTA	2309
Qy	2461	ATGCAGAGTTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC	2520
Db	2310	ATGCAGAGTTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC	2369
Qy	2521	TTCTATGATACAGGATACAGGACGCTTATATGGGTCACTCCCTGACCCAGATCAACAGGCG	2580
Db	2370	TTCTATGATACAGGATACAGGACGCTTATATGGGTCACTCCCTGACCCAGATCAACAGGCG	2429
Qy	2581	TATTACTTAGATCTGTGGCCATCCAGCAGAAAAGTTCCCTCTGACCAATCGTTTA	2640
Db	2430	TATTACTTAGATCTGTGGCCATCCAGCAGAAAAGTTCCCTCTGACCAATCGTTTA	2489
Qy	2641	CTGCTCTTACATGTTCTCGGATGAGAAATGTCATTTTGCACATACAGTATATATCTG	2700
Db	2490	CTGCTCTTACATGTTCTCGGATGAGAAATGTCATTTTGCACATACAGTATATATCTG	2549
Qy	2701	AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC	2760
Db	2550	AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC	2609
Qy	2761	AGCATAAGAGTTCTCTGAATCGGGAGAACATTTATGAATGTCATCTTTTGCACACTACCTTCAA	2820
Db	2610	AGCATAAGAGTTCTCTGAATCGGGAGAACATTTATGAATGTCATCTTTTGCACACTACCTTCAA	2669
Qy	2821	GAATAACCTTTGGATCAGCTATTGCTGCTTAAAGTGATATAATTTTGACCTGTGTAGAAC	2880
Db	2670	GAATAACCTTTGGATCAGCTATTGCTGCTTAAAGTGATATAATTTTGACCTGTGTAGAAC	2729
Qy	2881	TCCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA	2940
Db	2730	TCCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA	2789
Qy	2941	ATTGATCATCATTTTGATACCTGCGCATGTATCATCTCTCTGAAATTAATGTGGTG	3000
Db	2790	ATTGATCATCATTTTGATACCTGCGCATGTATCATCTCTCTGAAATTAATGTGGTG	2849
Qy	3001	CAATGAGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAACCCACATGCTCAAAA	3060
Db	2850	CAATGAGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAACCCACATGCTCAAAA	2909
Qy	3061	TCAAATGATACATATCTCTGAGAGACCCAGCATACCATTAAGATTAATTAATAAATAA	3120
Db	2910	TCAAATGATACATATCTCTGAGAGACCCAGCATACCATTAAGATTAATTAATAAATAA	2969
RESULT 12			
AY354202			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			
CDS			
Location/Qualifiers			
1..3030			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="15"			
/tissue_type="testis"			
1..3030			
/gene="Dpp8"			
161..2857			
/gene="Dpp8"			
/note="alternatively spliced"			
/codon_start=1			
/product="dipeptidyl peptidase 8 isoform 3"			
/protein_id="AA063887.1"			
/db_xref="GI:34329352"			
/translation="MKRSEQMKTKSGKCNMAAMETEQLGVETPETADEENIESQD			
RPKLPFVYVSWSLKKLLADTRKVKHYMMAKAPHDPMFVKRNDPDGPHSDRIYVL			
AMSGNRNTLYSEIKPTLNRAVLMSKPLLDLFOATLDYGMTSRBELLRKR			
TGTGIVSYDHYQSGTFLFQAGSGIVHVKGSGPQGTQOPLRPNLVETSCPNRMDP			
KLCPADPMIAFTHSNDIMLSNIVTREERLTVVHNELEMDASAGVATPVLOEE			
FNRGVYMKPKAETPSGKILALILEENDESEVELIHVTSPLMETRADSPRYPKT			
GTANPKVTFKSEIMIDAGRIIDVIOKELIQPELIFEGVEVIAAGWTPEGKYAW			
ILLDRSQRLQIVLISPELTIPVEDDMORORLIESVPDSVTPLIYEETDWINIH			
DIFVFPDQSHBEEIEFIFASECKTGFRHLKLTSLIKESKYKSSGLPASPDKPI			
KEEIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTQSDPLEHLYVYVYVNGEVTL			
TDGYSHSCCSIQHCDFFISKYSNQKPHCVSLYKSSPDEDDTCTKERNATILISA			
GLPDYTPPEIFSFESTTFTLQNLKPHDLQPKKIFIVLFIYGGPQQLVNNRPK			
GVKYPRLNLASLGVVVVYIDNRGSHRGKLFEGAFKMGQIEIDDVQVREGLOLASR			
YDFIDLRVGIHGWSYGSLSMALMORSDFRVAIAGAPVILWIFIDTGTIERIGH			
PDNQSGYILGSAVQAEKFPSENRLLHGLDENVHFATWISILLSLFLVRAGKPYD			
LQIYQERHSIRVPESGHEHYELHLLHYQLQENLGSRIAALKVI"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2806; Conservative			
89.9%; Score 2805.4; DB 8; Length 3030;			
99.9%; Pred. No. 0;			
0; Mismatches 1; Indels 0; Gaps 0;			
Qy	201	AGGAAATGCAACATGCGAGCAGCAATGGAACACAGACAGCTGGGTGTTGAGATATTGA	260
Db	196	AGGAAATGCAACATGCGAGCAGCAATGGAACACAGACAGCTGGGTGTTGAGATATTGA	255
Qy	261	AACCTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCTAAATTGGAGCCTTTTA	320
Db	256	AACCTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCTAAATTGGAGCCTTTTA	315
Qy	321	TGTTGAGCGGTATCTCTGAGTCAGCTTAAAAAGCTGCTGCCGATACCCAGAAAATATCA	380
Db	316	TGTTGAGCGGTATCTCTGAGTCAGCTTAAAAAGCTGCTGCCGATACCCAGAAAATATCA	375
Qy	381	TGGCTACATGATGGCTTAAGGCACCATGATTTTATGTTTCTGAAGAGGAATGATCCAGA	440
Db	376	TGGCTACATGATGGCTTAAGGCACCATGATTTTATGTTTCTGAAGAGGAATGATCCAGA	435

Qy 441 TGAACCTCATTCAGACAGAACTATTACCTTGCCATGTCTGGTGAGAACAGAGAAATATAC 500
Db 436 TGAACCTCATTCAGACAGAACTATTACCTTGCCATGTCTGGTGAGAACAGAGAAATATAC 495
Qy 501 ACTGTTTATCTGAAATTCCTCAAACTATCAATAGACGACGAGCTTAAATGCTCTCTTG 560
Db 496 ACTGTTTATCTGAAATTCCTCAAACTATCAATAGACGACGAGCTTAAATGCTCTCTTG 555
Qy 561 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAGAAGA 620
Db 556 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAGAAGA 615
Qy 621 AGAATATTAAAGAAAGAAACCGCATTTGGAACAGTGGAAATTCCTTCTTACGATATCA 680
Db 616 AGAATATTAAAGAAAGAAACCGCATTTGGAACAGTGGAAATTCCTTCTTACGATATCA 675
Qy 681 CCAAGCAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGAATTTATCACGTAAAGATGG 740
Db 676 CCAAGCAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGAATTTATCACGTAAAGATGG 735
Qy 741 AGGGCCACAAGGATTTACGCAACAACTTTAAAGGCCCAATCTAGTGGAACACTAGTTGTTC 800
Db 736 AGGGCCACAAGGATTTACGCAACAACTTTAAAGGCCCAATCTAGTGGAACACTAGTTGTTC 795
Qy 801 CAAATACCGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATACA 860
Db 796 CAAATACCGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATACA 855
Qy 861 TAGCAACGATATTTGGATATCTAAACATCGTAACCAAGAGAAAGGAGACTCACTTATGT 920
Db 856 TAGCAACGATATTTGGATATCTAAACATCGTAACCAAGAGAAAGGAGACTCACTTATGT 915
Qy 921 GCACATGAGCTAGCCCAATCGAAGAAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 980
Db 916 GCACATGAGCTAGCCCAATCGAAGAAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 975
Qy 981 TCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACAAC 1040
Db 976 TCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACAAC 1035
Qy 1041 TCCAGTGGTGGTAAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGA 1100
Db 1036 TCCAGTGGTGGTAAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGA 1095
Qy 1101 AATTATCATGTTACATCCCTATCTGTGMAACAGGAGGCGAGATTCATTCCGTTATCC 1160
Db 1096 AATTATCATGTTACATCCCTATCTGTGMAACAGGAGGCGAGATTCATTCCGTTATCC 1155
Qy 1161 TAAACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGC 1220
Db 1156 TAAACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGC 1215
Qy 1221 TGAAGGAAGGATCATAGATGTCATAGATAAGAACTAATTCAACTTTTGGAGATTCATTT 1280
Db 1216 TGAAGGAAGGATCATAGATGTCATAGATAAGAACTAATTCAACTTTTGGAGATTCATTT 1275
Qy 1281 TGAAGAGTTGAATATATTGCCAGAGCTGGATGAGTCTCTCAGGGAATAATGCTGGTC 1340
Db 1276 TGAAGAGTTGAATATATTGCCAGAGCTGGATGAGTCTCTCAGGGAATAATGCTGGTC 1335
Qy 1341 CATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCGAATATT 1400
Db 1336 CATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCGAATATT 1395
Qy 1401 TATCCCAAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTCAGTCAGTGCTGATTC 1460
Db 1396 TATCCCAAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTCAGTCAGTGCTGATTC 1455
Qy 1461 TGTGAGCCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACAT 1520
Db 1456 TGTGAGCCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACAT 1515
Qy 1521 CTTTCATGTTTTTCCCCAAAGTCACGAAGAGAAATTCAGTGTATTATTTTGGCTCTGAATG 1580

Db 1516 CTTTCATGTTTTTCCCCAAAGTCACGAAGAGAAATTCAGTGTATTATTTTGGCTCTGAATG 1575
Qy 1581 CAAAAACAGGTTTCCGTCATTATTATACAAAATTCATCTATTATTTAAAGGAAAGCAAAATATA 1640
Db 1576 CAAAAACAGGTTTCCGTCATTATTATACAAAATTCATCTATTATTTAAAGGAAAGCAAAATATA 1635
Qy 1641 ACATCCAGTGGTGGGCTGCCCTGCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGAT 1700
Db 1636 ACATCCAGTGGTGGGCTGCCCTGCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGAT 1695
Qy 1701 AGCAATTTACCAAGTGGTGAATGGAAGTCTTGCGCGCATCGATCTAATATCAAGTTGA 1760
Db 1696 AGCAATTTACCAAGTGGTGAATGGAAGTCTTGCGCGCATCGATCTAATATCAAGTTGA 1755
Qy 1761 TGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTGAAGCATCACCT 1820
Db 1756 TGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTGAAGCATCACCT 1815
Qy 1821 GTACGTAGTCAGTTACGTAAATCTTGGAAGGTGAACAAGGTGACTGACCGTGGCTACTC 1880
Db 1816 GTACGTAGTCAGTTACGTAAATCTTGGAAGGTGAACAAGGTGACTGACCGTGGCTACTC 1875
Qy 1881 ACATTTCTTGTGCATCAGTCAGACCTGTGACTCTTTTATTAAGTAAAGTATAGTAAACAGAA 1940
Db 1876 ACATTTCTTGTGCATCAGTCAGACCTGTGACTCTTTTATTAAGTAAAGTATAGTAAACAGAA 1935
Qy 1941 GAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAA 2000
Db 1936 GAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAA 1995
Qy 2001 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGAGGTCTCTTCTGACTATACTCC 2060
Db 1996 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGAGGTCTCTTCTGACTATACTCC 2055
Qy 2061 TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAA 2120
Db 2056 TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAA 2115
Qy 2121 GCCTCATGATCTACGCTGGAAAGAAATATCTACTGTCTGTGTTTATATATGGTGGTCC 2180
Db 2116 GCCTCATGATCTACGCTGGAAAGAAATATCTACTGTCTGTGTTTATATATGGTGGTCC 2175
Qy 2181 TCAGGTGTCAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 2240
Db 2176 TCAGGTGTCAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 2235
Qy 2241 AGCCTCTTAGTTTATGTGGTTGATGATAGAACAGGGGATCCTGTACCGAGGGCT 2300
Db 2236 AGCCTCTTAGTTTATGTGGTTGATGATAGAACAGGGGATCCTGTACCGAGGGCT 2295
Qy 2301 TAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGA 2360
Db 2296 TAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGA 2355
Qy 2361 AGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTTAGACTCGTGGGCGATCCA 2420
Db 2356 AGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTTAGACTCGTGGGCGATCCA 2415
Qy 2421 CGGTGTGCTTATGGAAGGATACCTCTCCCTGATGGCAATTAATGCAAGAGGTGATATCTT 2480
Db 2416 CGGTGTGCTTATGGAAGGATACCTCTCCCTGATGGCAATTAATGCAAGAGGTGATATCTT 2475
Qy 2481 CAGGTTGCTTATTTGCTGGGCGCCAGCTCACTCTGTGGATCTTCTATGATACAGGATACAC 2540
Db 2476 CAGGTTGCTTATTTGCTGGGCGCCAGCTCACTCTGTGGATCTTCTATGATACAGGATACAC 2535
Qy 2541 GGAACGTTATATGGGTCACTCCCTGACCAAGTGAACAGAGGCTATTACTTAGGATCTGTGGC 2600
Db 2536 GGAACGTTATATGGGTCACTCCCTGACCAAGTGAACAGAGGCTATTACTTAGGATCTGTGGC 2595
Qy 2601 CATCAAGCAAGAAAGTTCCCTCTGAAACAAATCGTTTACCTGCTCTTACATGGTTTCTT 2660

Db 2596 CATGCACGCAAAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCCT 2655
Qy 2661 GGATGAGAAATGTCATTTTGGCAATACACAGTATATTAATCTAGAGTTTCTTTAGTAGAGGCTGG 2720
Db 2656 GGATGAGAAATGTCATTTTGGCAATACACAGTATATTAATCTAGAGTTTCTTTAGTAGAGGCTGG 2715
Qy 2721 AAAGCCATPATGATTACAGATCTATCTCAGGAGACACAGATAAGATTCCTGAATC 2780
Db 2716 AAAGCCATPATGATTACAGATCTATCTCAGGAGACACAGATAAGATTCCTGAATC 2775
Qy 2781 GGGAGAAATATGAATGCACTCTTTTGGCACTACCTTTCAAGAAAACCTTGGATCACGAT 2840
Db 2776 GGGAGAAATATGAATGCACTCTTTTGGCACTACCTTTCAAGAAAACCTTGGATCACGAT 2835
Qy 2841 TGCTGCTCTTAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGTATACACTGGCTA 2900
Db 2836 TGCTGCTCTTAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGTATACACTGGCTA 2895
Qy 2901 TTTAACCAAAATGAGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCATTTTGTAT 2960
Db 2896 TTTAACCAAAATGAGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCATTTTGTAT 2955
Qy 2961 ACCTGCCATGTAAATCATCTACTCTCTGAAAATAAATGTGTGCCATGCA 3007
Db 2956 ACCTGCCATGTAAATCATCTACTCTCTGAAAATAAATGTGTGCCATGCA 3002

RESULT 13
LOCUS AR448400 2797 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1103 from patent US 6673549.
ACCESSION AR448400
VERSION AR448400.1 GI:42676724
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2797)
AUTHORS Furness,L.M. and Buchbinder,J.L.
TITLE Genes expressed in C3A liver cell cultures treated with steroids
JOURNAL Patent: US 6673549-A 1103 06-JAN-2004;
Incyte Corporation; Palo Alto, CA
FEATURES
source Location/Qualifiers
1..2797
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 88.4%; Score 2759.4; DB 6; Length 2797;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;
Qy 335 CCTGGAGTCAGCTTAAAGCTGCTGCCGATACCAGAAATATCATGGCTACATGATGG 394
Db 1 CCTGGAGTCAGCTTAAAGCTGCTGCCGATACCAGAAATATCATGGCTACATGATGG 60
Qy 395 CTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 454
Db 61 CTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 120
Qy 455 ACAGAACTATTACCTTCCCATGTCTGTGTAGAAACAGAGAAAATACACTGTTTATTCG 514
Db 121 ACAGAACTATTACCTTCCCATGTCTGTGTAGAAACAGAGAAAATACACTGTTTATTCG 180
Qy 515 AAATTCCTCCAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTGGAAAGCCCTTTTGG 574
Db 181 AAATTCCTCCAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTGGAAAGCCCTTTTGG 240
Qy 575 ATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAGAAGAAGAACTATTAAAG 634
Db 241 ATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAGAAGAAGAACTATTAAAG 300
Qy 635 AAAGAAAACGCAATTGGAACAGTCGGAATTTGCTTCTTACGATTATCAACCAAGGAAGTGAA 694

Db 301 AAAGAAAACGCAATTGGAACAGTCGGAATTTCTTCTTACGATTATCAACCAAGGAAGTGAA 360
Qy 695 CAATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGAGAGGCCACACAGGAT 754
Db 361 CAATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGAGAGGCCACACAGGAT 420
Qy 755 TTACCCCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTCTTCCCAACATACGATGG 814
Db 421 TTACCCCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTCTTCCCAACATACGATGG 480
Qy 815 ATCCAAAAATATGCCCCCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT 874
Db 481 ATCCAAAAATATGCCCCCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT 540
Qy 875 GGATATCTAAACATCGTAAACAGAGAAAGAGAGACTCACTTAATGTGCACAATGAGCTAG 934
Db 541 GGATATCTAAACATCGTAAACAGAGAAAGAGAGACTCACTTAATGTGCACAATGAGCTAG 600
Qy 935 CCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTTGTCTCCAAAGAAAT 994
Db 601 CCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTTGTCTCCAAAGAAAT 660
Qy 995 TTGATAGATATTCTGGCTATTGGTGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGGTA 1054
Db 661 TTGATAGATATTCTGGCTATTGGTGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGGTA 720
Qy 1055 AAATTCCTTAGAATTTCTATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTA 1114
Db 721 AAATTCCTTAGAATTTCTATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTA 780
Qy 1115 CATCCCTATGTTGGAACAAAGG-AGGCGAGATTCATTCCGTTATCTCTAAACAGGTGACA 1173
Db 781 CATCCCTATGTTGGAACAAAGGCAAGGCGAGATTCATTCCGTTATCTCTAAACAGGTGACA 840
Qy 1174 GCAATCTTAAAGTCATTTTAAAGTGCAGAAATATGATGATGATGCTGAGCAAGGATC 1233
Db 841 GCAATCTTAAAGTGCATTTTAAAGTGCAGAAATATGATGATGATGCTGAGCAAGGATC 900
Qy 1234 ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAGATTCATTTTGAAGAGTTGAA 1293
Db 901 ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAGATTCATTTTGAAGAGTTGAA 960
Qy 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATA- - -TGCTTGGTCCATCCTACT 1349
Db 961 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATA- - -TGCTTGGTCCATCCTACT 1020
Qy 1350 AGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGT 1409
Db 1021 AGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGT 1080
Qy 1410 AGAAGATGATGTTATGGAAGGCGAGAGACTCAITGAGTCAGTGCCTGATTTCTGTGAGCC 1469
Db 1081 AGAAGATGATGTTATGGAAGGCGAGAGACTCAITGAGTCAGTGCCTGATTTCTGTGAGCC 1140
Qy 1470 ACTAATTTCTATGAAGAAAACAAAGACATCTGGATAAATATCCATGACATCTTTTATGT 1529
Db 1141 ACTAATTTCTATGAAGAAAACAAAGACATCTGGATAAATATCCATGACATCTTTTATGT 1200
Qy 1530 TTTTCCCAAGTGCAGAGAGGAAATGAGTTTATTTTGGCTCTCAATGCAAAACAGG 1589
Db 1201 TTTTCCCAAGTGCAGAGAGGAAATGAGTTTATTTTGGCTCTCAATGCAAAACAGG 1260
Qy 1590 TTTCCGTCATTTATCAAAATTTACATCTATTTTAAAGGAAGAGAAATATAAAGATCCAG 1649
Db 1261 TTTCCGTCATTTATCAAAATTTACATCTATTTTAAAGGAAGAGAAATATAAAGATCCAG 1320
Qy 1650 TGGTGGGCTGCTGCTCCAAAGTGAATTCAGTGTCTTATCAAGAGAGATAGCAATTC 1709
Db 1321 TGGTGGGCTGCTGCTCCAAAGTGAATTCAGTGTCTTATCAAGAGAGATAGCAATTC 1380
Qy 1710 CAGTGTGAATGGGAAGTTTGTGGCGGCAATGGAATCTAATATCCAAAGTGTGATGAAGTCAG 1769

Db 1381 CAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAAATATCCAAAGTTGATGAAGTCAG 1440
Qy 1770 AAGGCTGGTATATTTTGAAGGCAACAAGACTCCCTTTAGAGCATCACTGTACGTAGT 1829
Db 1441 AAGGCTGGTATATTTTGAAGGCAACAAGACTCCCTTTAGAGCATCACTGTACGTAGT 1500
Qy 1830 CAGTTACGTAAATCTCGAGAGGTGACAAAGCTGACGTACGCGGTACTCACATTTCTTG 1889
Db 1501 CAGTTACGTAAATCTCGAGAGGTGACAAAGCTGACGTACGCGGTACTCACATTTCTTG 1560
Qy 1890 CTGCAATCAGTCAGCACTGTGACTCTCTTTATAAGTAAGTATAGTAACACAGAGAATCCACA 1949
Db 1561 CTGCAATCAGTCAGCACTGTGACTCTCTTTATAAGTAAGTATAGTAACACAGAGAATCCACA 1620
Qy 1950 CTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAAGGA 2009
Db 1621 CTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAAGGA 1680
Qy 2010 ATTTTGGGCGACCATTTTGGATTTGCAATTCAGAGGTCCTCTTCTGACTATCTCTCCAGAAAT 2069
Db 1681 ATTTTGGGCGACCATTTTGGATTTGCAATTCAGAGGTCCTCTTCTGACTATCTCTCCAGAAAT 1740
Qy 2070 TTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCCTCATGA 2129
Db 1741 TTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCCTCATGA 1800
Qy 2130 TCTACAGCTCGAAGAAATATCTTACTGTGCTGCTTCATATATGATGGTCTCTCAGGTGCA 2189
Db 1801 TCTACAGCTCGAAGAAATATCTTACTGTGCTGCTTCATATATGATGGTCTCTCAGGTGCA 1860
Qy 2190 GTTGTGTAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCAAGCTCTCT 2249
Db 1861 GTTGTGTAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCAAGCTCTCT 1920
Qy 2250 AGTTATGTGGTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA 2309
Db 1921 AGTTATGTGGTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA 1980
Qy 2310 AGCGGCTTTAAATATAAAATGGGTCAAAATGAAGATTTGACATCAGGTGAAGGACTCCA 2369
Db 1981 AGCGGCTTTAAATATAAAATGGGTCAAAATGAAGATTTGACATCAGGTGAAGGACTCCA 2040
Qy 2370 ATATCTAGCTTCTCGATATGATTTCAATTTGACTTGAATCTGTGGGATCCACGGCTGGTC 2429
Db 2041 ATATCTAGCTTCTCGATATGATTTCAATTTGACTTGAATCTGTGGGATCCACGGCTGGTC 2100
Qy 2430 CTATGAGGATACCTCTCCCTGATGGCAATTAATCAGAGGTCAAGATATCTTCAGGGTTCG 2489
Db 2101 CTATGAGGATACCTCTCCCTGATGGCAATTAATCAGAGGTCAAGATATCTTCAGGGTTCG 2160
Qy 2490 TATTGCTGGGGCCCGAGTCACCTCTGTGGATCTCTTATGATACAGGATACACGGAACGTGA 2549
Db 2161 TATTGCTGGGGCCCGAGTCACCTCTGTGGATCTCTTATGATACAGGATACACGGAACGTGA 2220
Qy 2550 TATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGC 2609
Db 2221 TATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGC 2280
Qy 2610 AGAAAAGTTCCCTCTGCAACCAATCGTTTACTGCTTTACATGTTTCTCGATGAGAA 2669
Db 2281 AGAAAAGTTCCCTCTGCAACCAATCGTTTACTGCTTTACATGTTTCTCGATGAGAA 2340
Qy 2670 TGTCCATTTTGCACATACAGTATATTAAGTATTTTATGAGGCTGGAAAGCCATA 2729
Db 2341 TGTCCATTTTGCACATACAGTATATTAAGTATTTTATGAGGCTGGAAAGCCATA 2400
Qy 2730 TGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTCTCTGATTCGGGAGACA 2789
Db 2401 TGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTCTCTGATTCGGGAGACA 2460
Qy 2790 TTATGAACTGCACTTTTGGCACTACTCTTCAGAAAACCTTGGATCAGTATTCCTGCTCT 2849
Db 2461 TTATGAACTGCACTTTTGGCACTACTCTTCAGAAAACCTTGGATCAGTATTCCTGCTCT 2520

Qy 2850 AAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTATAACCAA 2909
Db 2521 AAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTATAACCAA 2580
Qy 2910 ATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATTTTGTGATCTGCGCAT 2969
Db 2581 ATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATTTTGTGATCTGCGCAT 2640
Qy 2970 GTPAACATCTACTCTCGAAAAATAATGTGTGCCATGACAGGGGTCTACGGTTTGTGGTAGT 3029
Db 2641 GTPAACATCTACTCTCGAAAAATAATGTGTGCCATGACAGGGGTCTACGGTTTGTGGTAGT 2700
Qy 3030 AATCTTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCA 3089
Db 2701 AATCTTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCA 2760
Qy 3090 GCAATACCAATAAGAAATTACTTAAAAAATAAAAA 3120
Db 2761 GCAATACCAATAAGAAATTACTTAAAAAATAAAAA 2791

RESULT 14

BC030688 3125 bp mRNA linear PRI 28-JUL-2005
DEFINITION Homo sapiens dipeptidylpeptidase 8, transcript variant 4, mRNA
LOCUS (cDNA clone MGC:26191 IMAGE:4822550), complete cds.
ACCESSION BC030688
VERSION BC030688.2 GI:34190028
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3125)
AUTHORS Klausner, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Krausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toohy, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Adams, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3125)
NTH MGC Project
Direct Submission
Submitted (24-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug '25, 2003 this sequence version replaced gi:21265132.
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

http://www.systemsbio.org

contact: amand@systemsbio.org

Anup Madan, Jessica Fahy, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 33 Row: d Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 37577092.

FEATURES

source

1..3125
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:26191 IMAGE:4822550"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="vector: pBluescriptR"

gene

1..3125
/gene="DPP8"
/note="synonyms: FLJ20283, FLJ14920, MGC26191, MSTP141"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"

CDS

319..2862
/gene="DPP8"
/codon_start=1
/product="dipeptidyl peptidase 8, isoform 4"
/protein_id="AAH30688.3"
/db_xref="GI:54114905"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"
/translation="MWKRSQMKIKSGKCNMAAMETEOLGVEIFETADCBENIESQD
RPKLEPFVVERYSQKLLADTRKYHGVMKAPDHFVFRKNDPDGPHSDRIYYL
AMSGENRNTLFYSEIPTKINRAAIVLMKPLDLFQATLDYGMYSREELLRRERK
IGTVIASDYHGGSGTFLFQAGSGIYHVKGQPGQFTQOPLRPNLVETSPNIRMDP
KLCPADPMWAFIHNDIWI SNITVREERLTYVHNELANMEEDARSAGVATFVLQEE
FDYSGYVMCWPKEASTPSSGKILRIYENDESEVEI IHVTSPLMTRRADSPRYPKT
GTAMPKVTFKMEIMDAEGRIIDVIDKELIQPEILFEGVEYIARAQWTEPGKYAMS
ILLDRSTRLIQIVLISPELIPVEDDVMERQRLIESVPSVTPLIIEYETDIDINIH
DI FHVFPQHEEETEFIPASECTGFHRLVYVPGTKDSPLEHHLVGVVNPGEVTRL
KEEIALTSGHEVLGRHSGNIQVDEVRLLVYVPGTKDSPLEHHLVGVVNPGEVTRL
TDGYSHSCCI SOHCDPFI SKYSNKNPHCVSLYKLSPPEDDPTCKTKEFWATILDSA
GPLDYPPEI FSPESSTGTLYGMLYKPHDLPQKKYPTLVFIYGGQVOLVNNRPK
GVKYFRNLTLASLGYVVIDNRSGCHRLKPEGAPKYRWAIAGAPVTLMTFVDTGY
TERYMGHPDNEQGYLGSVAMQAEKFPSEPNELLLHGLDENVHFATSLILLSFLV
RAGKPYLIQITPQERHSIRVPESGEHYELHLLHYLQENLGSRIALKVI"

ORIGIN

Query Match 85.1%; Score 2656.4; DB 8; Length 3125;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 2966; Conservative 0; Mismatches 1; Indels 289; Gaps 2;
1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGTGTACTTGC CGCGCGCTTCTTAGTGCGG 60
18 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGTGTACTTGC CGCGCGCTTCTTAGTGCGG 77
61 CGTTCCGCGCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCACTGCAACACGAGACCG 120
78 CGTTCCGCGCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCACTGCAACACGAGACCG 137
121 GAGTGAGGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGGCGCAGCTCGGAGCG 180
138 GAGTGAGGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGGCGCAGCTCGGAGCG 197
181 TCCGGCGCGCGCGGGGGA----- 200
198 TCCGGCGCGCGCGGGGGAGGCTTTGATGGTGAAGAAAGGAAAGATTCTTGAGAGAG 257
201 ----- 200

258 AGCAGGATGACGAGGGGATTCTATGCTTGAAGTCGAGTCACCTTGAAGAAAGATCTCTTTG 317
201 -----AGGAAATGCAACATGCGCAGCAGC 224
318 AATGTGGAAGAGATCTGACGAGATGAAATAAATAATCAGGAAATGCAACATGCGCAGCAGC 377
225 AATGAAACAGAACAGCTGGGTGTGAGATATTTGAACCTCGGACTGTGAGGAGAAATAT 284
378 AATGAAACAGAACAGCTGGGTGTGAGATATTTGAACCTCGGACTGTGAGGAGAAATAT 437
285 TGAATCAGAGATCGGCTAAATTTGAGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCA 344
438 TGAATCAGAGATCGGCTAAATTTGAGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCA 497
345 GCTTAAAAAGCTGCTTGCAGATACCAAGAAATATCATGGCTACATGATGGCTAAGGACCC 404
498 GCTTAAAAAGCTGCTTGCAGATACCAAGAAATATCATGGCTACATGATGGCTAAGGACCC 557
405 ACATGATTTTCATGTTTGTGGAAGAGAAATGATCCAGATGGACCTCATTCAGACAGAAATCTA 464
558 ACATGATTTTCATGTTTGTGGAAGAGAAATGATCCAGATGGACCTCATTCAGACAGAAATCTA 617
465 TTACCTTGGCATGCTCGGTGAGAAACAGAGAAATATACACTGTTTATTTCTGAAATTCCTCAA 524
618 TTACCTTGGCATGCTCGGTGAGAAACAGAGAAATATACACTGTTTATTTCTGAAATTCCTCAA 677
525 AACTATCAATAGACGAGCAGTCTTAATGCTCTCTTGAAGCCCTTTTGGATCTTTTCA 584
678 AACTATCAATAGACGAGCAGTCTTAATGCTCTCTTGAAGCCCTTTTGGATCTTTTCA 737
585 GGCACACTGGAATGTAATCTCGAGAGAGAACTATTAGAGAGAGAAAGAACG 644
738 GGCACACTGGAATGTAATCTCGAGAGAGAACTATTAGAGAGAGAAAGAACG 797
645 CATTTGAAACAGTCGGAATGCTTCTTAGATTTACCAAGAGAGAGTGAACATTTCTGTT 704
798 CATTTGAAACAGTCGGAATGCTTCTTAGATTTACCAAGAGAGTGAACATTTCTGTT 857
705 TCAAGCCGCTAGTGGAAATTTATACGTAAGAGATGAGAGGCGCAACAGGATTTACGCAACA 764
858 TCAAGCCGCTAGTGGAAATTTATACGTAAGAGATGAGAGGCGCAACAGGATTTACGCAACA 917
765 ACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTTGCCAAACATACGAGATGATCCAAAT 824
918 ACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTTGCCAAACATACGAGATGATCCAAAT 977
825 ATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAA 884
978 ATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAA 1037
885 CATCGTAACCCAGAGAGAAAGAGACTCACCTTATGTGCACATGAGCTAGCACAACATGGA 944
1038 CATCGTAACCCAGAGAGAAAGAGACTCACCTTATGTGCACATGAGCTAGCACAACATGGA 1097
945 AGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCCAAGAGAAATTTGATAGATA 1004
1098 AGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCAGAGAAATTTGATAGATA 1157
1005 TTCTGCTATTGGTGTGTCAAAAGCTGAAACAACTCCGAGTGGTGGTAAATTTCTTAG 1064
1158 TTCTGCTATTGGTGTGTCAAAAGCTGAAACAACTCCGAGTGGTGGTAAATTTCTTAG 1217
1065 AATTCATATGAAGAAATATGATGAATCTGAGGTGAAATTTATCATGTTATCATCCCTAT 1124
1218 AATTCATATGAAGAAATATGATGAATCTGAGGTGAAATTTATCATGTTATCATCCCTAT 1277
1125 GTTGGAAACAGAGAGGCGCAGATTCATTCGGTTATCTTAAAAACAGGTACAGAAATTCCTAA 1184
1278 GTTGGAAACAGAGAGGCGCAGATTCATTCGGTTATCTTAAAAACAGGTACAGAAATTCCTAA 1337
1185 AGTCACCTTTAAGATGTCAGAAATATGATGATGCTGAAGAGAGGATCATAGATGTCAT 1244
1338 AGTCACCTTTAAGATGTCAGAAATATGATGATGCTGAAGAGAGGATCATAGATGTCAT 1397

1245 AGATAAGGAACAAATTTCAACCTTTTGAGATTTCTATTTTGAAGAGTTGAATATATTTGCCAG 1304
1398 AGATAAGGAACAAATTTCAACCTTTTGAGATTTCTATTTTGAAGAGTTGAATATATTTGCCAG 1457
1305 AGCTGGATGACCTCTGAGGGAATAATATCTTTGGTCCATCTCTACTAGATCGCTCCAGAC 1364
1458 AGCTGGATGACCTCTGAGGGAATAATATCTTTGGTCCATCTCTACTAGATCGCTCCAGAC 1517
1365 TCGCCTACAGATAGTTGATCTCACCTCGAATATTTATTTATCCAGTAGAAGATGATGTTAT 1424
1518 TCGCCTACAGATAGTTGATCTCACCTCGAATATTTATTTATCCAGTAGAAGATGATGTTAT 1577
1425 GGAAGGACAGAGCTCATTTAGAGTCAAGTCTGATCTCTGAGCCGACCTAATATATATGA 1484
1578 GGAAGGACAGAGCTCATTTAGAGTCAAGTCTGATCTCTGAGCCGACCTAATATATATGA 1637
1485 AGAAACAAACAGATCTGGATAAATATCCATGACATCTTTTATGTTTTTCCCAAGTCA 1544
1638 AGAAACAAACAGATCTGGATAAATATCCATGACATCTTTTATGTTTTTCCCAAGTCA 1697
1545 CGAAGGAATTTGAGTTATTTTGGCTCTCAATGCAAAACAGGTTTTCGTCAATTTATA 1604
1698 CGAAGGAATTTGAGTTATTTTGGCTCTCAATGCAAAACAGGTTTTCGTCAATTTATA 1757
1605 CAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGGCTGCCTGC 1664
1758 CAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGGCTGCCTGC 1817
1665 TCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGA 1724
1818 TCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGA 1877
1725 AGTTCTTGGCGGATGGATCTAATATCAAGTGGATGAAGTGAAGAGCTGGTGAATTTT 1784
1878 AGTTCTTGGCGGATGGATCTAATATCAAGTGGATGAAGTGAAGAGCTGGTGAATTTT 1937
1785 TGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTAGCTAGTACGTAAATTC 1844
1938 TGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTAGCTAGTACGTAAATTC 1997
1845 TGGAGAGTGACAAAGGCTGACCGTGGCTACTCAATTTCTGCTGATCAGTCAGCA 1904
1998 TGGAGAGTGACAAAGGCTGACCGTGGCTACTCAATTTCTGCTGATCAGTCAGCA 2057
1905 CTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTA 1964
2058 CTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTA 2117
1965 CAAGCTATCAAGTCTCAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGGCAACAT 2024
2118 CAAGCTATCAAGTCTCAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGGCAACAT 2177
2025 TTTGGATTCAGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAG 2084
2178 TTTGGATTCAGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAG 2237
2085 TACTACTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAA 2144
2238 TACTACTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAA 2297
2145 GAAATATCTACTGTCTGTTTCAATATATGGTGGTCTCTAGGTGCAAGTGGTGAATATCG 2204
2298 GAAATATCTACTGTCTGTTTCAATATATGGTGGTCTCTAGGTGCAAGTGGTGAATATCG 2357
2205 GTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTT 2264
2358 GTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTT 2417
2265 AGTGATGACAAACAGGGGATCTCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATA 2324
2418 AGTGATGACAAACAGGGGATCTCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATA 2477

2325 TAAAATGGGTCAAATAGAAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCG 2384
2478 TAAAAT----- 2483
2385 ATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCT 2444
2484 ----- 2483
2445 CTCCCTGATGGCATTAATGCGAGGCTCAGATATCTTCAAGGCTTGCTATTGCTGGGGCCCC 2504
2484 ----- 2504
2505 AGTCACCTCTGTGGATCTTCTATGATACAGGATACACGAAAGTGTATATGGGTCAACCTGA 2564
2505 AGTCACCTCTGTGGATCTTCTATGATACAGGATACACGAAAGTGTATATGGGTCAACCTGA 2564
2565 CCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTC 2624
2565 CCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTC 2624
2625 TGAACCAAAATCGTTTACTCTCTTACATGTTTCTGATGAGAAATGTCATTTTGACACA 2684
2625 TGAACCAAAATCGTTTACTCTCTTACATGTTTCTGATGAGAAATGTCATTTTGACACA 2684
2685 TACCAGTATATTAATCTGAGTTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTA 2744
2685 TACCAGTATATTAATCTGAGTTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTA 2744
2745 TCCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATTTATGAATGATATCT 2804
2745 TCCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATTTATGAATGATATCT 2804
2805 TTTGCACTACCTTCAAGAAAACCTTGATCAGTATTCCTCTTAAAGTGATATAAT 2864
2805 TTTGCACTACCTTCAAGAAAACCTTGATCAGTATTCCTCTTAAAGTGATATAAT 2864
2865 TTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTATAACCAATGAGGAGTTTAAAT 2924
2865 TTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTATAACCAATGAGGAGTTTAAAT 2924
2925 CAAACAGAAAACACAGAAATGATCATCACATTTTGTATCTGCCATGTAACTACTACTCT 2984
2925 CAAACAGAAAACACAGAAATGATCATCACATTTTGTATCTGCCATGTAACTACTACTCT 2984
2985 GAAATATAATGTGGTGCCATGCAAGGGTCTACGGTTTGTGGTGTAGTAACTTAATACCTTAA 3044
2985 GAAATATAATGTGGTGCCATGCAAGGGTCTACGGTTTGTGGTGTAGTAACTTAATACCTTAA 3044
3045 CCCACATGCTCAAAATCAAAATGATATATCTCTGAGAGACCCAGCAATACCATAGAA 3104
3045 CCCACATGCTCAAAATCAAAATGATATATCTCTGAGAGACCCAGCAATACCATAGAA 3104
3105 TTACTAAAAA 3120
3105 TTACTAAAAA 3120

RESULT 15
AX354795
LOCUS AX354795 2649 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 3 from Patent WO0179473.
ACCESSION AX354795
VERSION AX354795.1 GI:18619528
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Meyers, R.A. and Williamson, M.
TITLE 21553, a human prolyl oligopeptidase family member and uses thereof
JOURNAL Patent: WO 0179473-A 3 25-OCT-2001;

FEATURES		Millennium Pharmaceuticals, Inc. (US)	
source	1...2649	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match		84.8%; Score 2645.8; DB 6; Length 2649;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 2647; Conservative		0; Mismatches 2; Indels 0; Gaps 0;	
QY	214	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT	273
DB	1	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT	60
QY	274	GAGGAGAATATTTGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT	333
DB	61	GAGGAGAATATTTGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT	120
QY	334	TCCTGGAGTCAGCTTAAAGAGCTGCTTGCAGATACAGAGAAATATATGCTACATGATG	393
DB	121	TCCTGGAGTCAGCTTAAAGAGCTGCTTGCAGATACAGAGAAATATATGCTACATGATG	180
QY	394	GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	453
DB	181	GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	240
QY	454	GACAGAACTTATACCTTTGCCATGCTCTGGTGAAGACAGAGAAATAACACTGTTTATTTCT	513
DB	241	GACAGAACTTATACCTTTGCCATGCTCTGGTGAAGACAGAGAAATAACACTGTTTATTTCT	300
QY	514	GAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTG	573
DB	301	GAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTG	360
QY	574	GATCTTTTCAGGCAACACTGGAATGGAATGTTATTCGAGAAGAGAACTATTTAAGA	633
DB	361	GATCTTTTCAGGCAACACTGGAATGGAATGTTATTCGAGAAGAGAACTATTTAAGA	420
QY	634	GAAGAAAACGCAATTTGGAACAGTCGGAAATGCTTTTACGATATATCACAAAGGAAGTGA	693
DB	421	GAAGAAAACGCAATTTGGAACAGTCGGAAATGCTTTTACGATATATCACAAAGGAAGTGA	480
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAGATGGAGGGCCACAAAGA	753
DB	481	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAGATGGAGGGCCACAAAGA	540
QY	754	TTTACGCAACAAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGGATG	813
DB	541	TTTACGCAACAAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGGATG	600
QY	814	GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGTTTATACATAGCAACGATATT	873
DB	601	GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGTTTATACATAGCAACGATATT	660
QY	874	TGGATATCTAAATCGTAAACAGAGAAGAAAGGAGACTCACTTATGTGCACAATAGCTA	933
DB	661	TGGATATCTAAATCGTAAACAGAGAAGAAAGGAGACTCACTTATGTGCACAATAGCTA	720
QY	934	GCACAATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA	993
DB	721	GCACAATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA	780
QY	994	TTTGATAGATATTCCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGT	1053
DB	781	TTTGATAGATATTCCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGT	840
QY	1054	AAAAATCTTTAGAAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113
DB	841	AAAAATCTTTAGAAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	900
QY	1114	ACATCCCTCTATGTTGGAAACAGGAGGGCAGATTCATTCGGTTATCCTTAAACACAGGTACA	1173

901	ACATCCCTTATGTTGAAACAGGAGGCGCAGATTCATTCCTGTTATCTTAAACAGGTACA	960
1174	GCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGAGGAGGATC	1233
961	GCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGAGGAGGATC	1020
1234	ATAGATGTCATAGATAAGGAACTAAATCAACTTTTGGAGATCTTATTTTGAAGGAGTTGAA	1293
1021	ATAGATGTCATAGATAAGGAACTAAATCAACTTTTGGAGATCTTATTTTGAAGGAGTTGAA	1080
1294	TATATTTGCCAGAGCTCGATGGACTCTCTGAGGGAAAAATATGCTTGGTGCATCTCTACTAGAT	1353
1081	TATATTTGCCAGAGCTCGATGGACTCTCTGAGGGAAAAATATGCTTGGTGCATCTCTACTAGAT	1140
1354	CGGTCCCGACACTCGCCTACAGATAGTGTGATCTCACTCGAATATATTTATCCAGTAGAA	1413
1141	CGCTCCCGACACTCGCCTGCAGATAGTGTGATCTCACTCGAATATATTTATCCAGTAGAA	1200
1414	GATGATGTTATGAAAAGGCGAGAGACTCATTGGAGTCAGTGCCCTGATCTTGTGACGCGCACTA	1473
1201	GATGATGTTATGAAAAGGCGAGAGACTCATTGGAGTCAGTGCCCTGATCTTGTGACGCGCACTA	1260
1474	ATTATCTATGAAAGAAACAAACAGACATCTCGATAAATATCCATGACATCTTTTCAATGTTTTT	1533
1261	ATTATCTATGAAAGAAACAAACAGACATCTCGATAAATATCCATGACATCTTTTCAATGTTTTT	1320
1534	CCCCAAAGTCAACGAAGAGGAAATTTGAGTTTATTTTTTGCCTCTGAAATGCAAAAACAGGTTTC	1593
1321	CCCCAAAGTCAACGAAGAGGAAATTTGAGTTTATTTTTTGCCTCTGAAATGCAAAAACAGGTTTC	1380
1594	CGTCATTTATACAAAAATTTACATCTATTTTTTAAAGGAAAGCAAAATATAAAACGATCCAGTGGT	1653
1381	CGTCATTTATACAAAAATTTACATCTATTTTTTAAAGGAAAGCAAAATATAAAACGATCCAGTGGT	1440
1654	GGGTGCTGCTGCCAAGTGATTTTCAAGTGTCCTATCAAAAGAGGAGATAGCAATATACCAGT	1713
1441	GGGTGCTGCTGCCAAGTGATTTTCAAGTGTCCTATCAAAAGAGGAGATAGCAATATACCAGT	1500
1714	GGTGAATGGGAAGTCTTTTGGCCGCGCATGGATCTAAATATCCAAGTTGATCAAGTTCAGAAAG	1773
1501	GGTGAATGGGAAGTCTTTTGGCCGCGCATGGATCTAAATATCCAAGTTGATCAAGTTCAGAAAG	1560
1774	CTGGTATATTTTGAAGGCAACCAAGACTCCCCCTTTTAGAGCATCACCTGTACGTAGTCAAGT	1833
1561	CTGGTATATTTTGAAGGCAACCAAGACTCCCCCTTTTAGAGCATCACCTGTACGTAGTCAAGT	1620
1834	TACGTAATCTCGAGAGGTGCAAGAGCTGCATGACCGTGGCTACTCACATTTCTTCTCTGC	1893
1621	TACGTAATCTCGAGAGGTGCAAGAGCTGCATGACCGTGGCTACTCACATTTCTTCTCTGC	1680
1894	ATCAGTCAGCATGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAGAAATCCACACTGT	1953
1681	ATCAGTCAGCATGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAGAAATCCACACTGT	1740
1954	GTGTCCTTTTCAAGCTATCAAGTCTCTGAGATGACCCCACTTGCAGAAAACAAAGGAATTTT	2013
1741	GTGTCCTTTTCAAGCTATCAAGTCTCTGAGATGACCCCACTTGCAGAAAACAAAGGAATTTT	1800
2014	TGGGCCACCATTTTGGATTCAGCAGGTCTCTCTGACTATACTCTCCAGAAATTTTTC	2073
1801	TGGGCCACCATTTTGGATTCAGCAGGTCTCTCTGACTATACTCTCCAGAAATTTTTC	1860
2074	TCTTTTGAAGTACTACTCGGATTTTACATGTATGGGATGCTCTTCAAGAGCTCATGATCTA	2133
1861	TCTTTTGAAGTACTACTCGGATTTTACATGTATGGGATGCTCTTCAAGAGCTCATGATCTA	1920
2134	CAGCCTGGAAAAGAAATATCTCTACTGCTGTTTCATATATGTTGGTCTCTCAGGTGCAAGT	2193
1921	CAGCCTGGAAAAGAAATATCTCTACTGCTGTTTCATATATGTTGGTCTCTCAGGTGCAAGT	1980
2194	GTCAATAATTCGGTTTTTAAAGGAGTCAAGTATTTTCGGCTTGAATAACCTTAGCCTCTCTAGGT	2253

901	ACATCCCTATGTTGGAAACAGGAGGGCAGATTCAATTCGGTTATCCTTAAACACAGGTACA	960
1174	GCAATCTTAAGTCACTCTTTAAGATGTCAGAAATAATGATTGATGCTGGAAGGAAGATC	1233
961	GCAATCTTAAGTCACTCTTTAAGATGTCAGAAATAATGATTGATGCTGGAAGGAAGATC	1020
1234	ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAA	1293
1021	ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAA	1080
1294	TATATTTCCAGAGCTGGATGGAATCTCTGAGGGAATAATGCTTGGCTCCTACTACTAGAT	1353
1081	TATATTTCCAGAGCTGGATGGAATCTCTGAGGGAATAATGCTTGGCTCCTACTACTAGAT	1140
1354	CGCTCCCAAGCTCGCCCTACAGATAGTGTGATCTCACTGAAATATTTATCCCACTAGAA	1413
1141	CGCTCCCAAGCTCGCCCTACAGATAGTGTGATCTCACTGAAATATTTATCCCACTAGAA	1200
1414	GATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCACTGCTGATCTGTGTGACGCCACTA	1473
1201	GATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCACTGCTGATCTGTGTGACGCCACTA	1260
1474	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCACTGTTT	1533
1261	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCACTGTTT	1320
1534	CCCCAAAGCTCAGGAAGAGGAAATTTGAGTTTATTTTGGCTCTGGAATGCAAAAACAGGTTTC	1593
1321	CCCCAAAGCTCAGGAAGAGGAAATTTGAGTTTATTTTGGCTCTGGAATGCAAAAACAGGTTTC	1380
1594	CGTCATTTATCAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAACGATCCAGTGT	1653
1381	CGTCATTTATCAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAACGATCCAGTGT	1440
1654	GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCCTATCAAGAGGAGATAGCAATTTACCACT	1713
1441	GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCCTATCAAGAGGAGATAGCAATTTACCACT	1500
1714	GGTGAATGGGAAGTTCTTTGGCCGCGATGGATCTAATATCCAACTGATGAAGTCAGAAG	1773
1501	GGTGAATGGGAAGTTCTTTGGCCGCGATGGATCTAATATCCAACTGATGAAGTCAGAAG	1560
1774	CTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGGATCACTGATCTTCTGCTGC	1833
1561	CTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGGATCACTGATCTTCTGCTGC	1620
1834	TACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAATTTCTGCTGC	1893
1621	TACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAATTTCTGCTGC	1680
1894	ATCAGTCAGCACTGTGACTTTCTTTATAAGTAAGTATAGTAAACAGAGAATCCCACTGT	1953
1681	ATCAGTCAGCACTGTGACTTTCTTTATAAGTAAGTATAGTAAACAGAGAATCCCACTGT	1740
1954	GTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAAACAAAGGAATTT	2013
1741	GTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAAACAAAGGAATTT	1800
2014	TGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGCTGACTACTACTCTCCAGAAAATTTTC	2073
1801	TGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGCTGACTACTACTCTCCAGAAAATTTTC	1860
2074	TCCTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGACCTCATGATCTA	2133
1861	TCCTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGACCTCATGATCTA	1920
2134	CAGCTGGAAGAAATATCTACTGCTGCTTTTATATATGTTGGTCTCTCAGGTGAGTTG	2193
1921	CAGCTGGAAGAAATATCTACTGCTGCTTTTATATATGTTGGTCTCTCAGGTGAGTTG	1980
2194	GTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATCCCTAGCCTCTCTAGGT	2253

Db 1981 GTGAATAATCGGTTAAAGAGTCAAGTATTTCCGGTTGAAATACCCCTAGCCTCTCTAGGT 2040
QY 2254 TATGTGTTGTAGTAGACACACAGGGGATCCTGTCAACCGAGGCTTAAATTTGAAGGC 2313
Db 2041 TATGTGTTGTAGTAGACACACAGGGGATCCTGTCAACCGAGGCTTAAATTTGAAGGC 2100
QY 2314 GCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGAAGGACTCCAATAT 2373
Db 2101 GCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGAAGGACTCCAATAT 2160
QY 2374 CTAGCTTCTCGATATGATTTTCATTGACTTTAGATCGTGTGGCATCCACGGCTGCTCTAT 2433
Db 2161 CTAGCTTCTCGATATGATTTTCATTGACTTTAGATCGTGTGGCATCCACGGCTGCTCTAT 2220
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGGTTGCTATT 2493
Db 2221 GGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGGTTGCTATT 2280
QY 2494 GCTGGGGCCCGAGTCACCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2553
Db 2281 GCTGGGGCCCGAGTCACCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2340
QY 2554 GGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAA 2613
Db 2341 GGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAA 2400
QY 2614 AAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTTACATGGTTTCTCTGGATGAGATGTC 2673
Db 2401 AAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTTACATGGTTTCTCTGGATGAGATGTC 2460
QY 2674 CATTTTGCACATACCAAGTATATTACTGAGTTTTTTTGTAGGGCTCGAAAGCCATATGAT 2733
Db 2461 CATTTTGCACATACCAAGTATATTACTGAGTTTTTTTGTAGGGCTCGAAAGCCATATGAT 2520
QY 2734 TTACAGATCTATCTCAGGAGAGACACAGCATTAAGATTCCTGAATCGGGAGACATTTAT 2793
Db 2521 TTACAGATCTATCTCAGGAGAGACACAGCATTAAGATTCCTGAATCGGGAGACATTTAT 2580
QY 2794 GAACCTGCATCTTTTGCATCTTCAAGAAAACCTTGGATCAGTATTCCTCTCTAAAA 2853
Db 2581 GAACCTGCATCTTTTGCATCTTCAAGAAAACCTTGGATCAGTATTCCTCTCTAAAA 2640
QY 2854 GTGATATAA 2862
Db 2641 GTGATATAA 2649

Search completed: April 15, 2006, 00:03:00
Job time : 15361 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 19:20:55 ; Search time 1768 Seconds
(without alignments)
11761.222 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgctaaagctccgag.....agaattactaaaaaaaaa 3120

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: geneseq1980s.*
- 2: geneseq1990s.*
- 3: geneseq2000s.*
- 4: geneseq2001as.*
- 5: geneseq2001bs.*
- 6: geneseq2002as.*
- 7: geneseq2002bs.*
- 8: geneseq2003as.*
- 9: geneseq2003bs.*
- 10: geneseq2003cs.*
- 11: geneseq2003ds.*
- 12: geneseq2004as.*
- 13: geneseq2004bs.*
- 14: geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3120	100.0	3120	4 AAC85694	Aac85694 Nucleotid
2	3120	100.0	3120	6 AAD38956	Aad38956 Human dip
3	3106.4	99.6	4829	6 ABK83327	Abk83327 cDNA enco
4	3105.2	99.5	3143	6 AAH99934	Aah99934 cDNA enco
5	3101.8	99.4	3106	6 ABK12892	Abk12892 Human pro
6	2824.4	90.5	4685	6 ABK83332	Abk83332 cDNA enco
7	2806.4	89.9	4676	6 ABK83331	Abk83331 cDNA enco
8	2785.8	89.3	2952	10 ACA92421	Ac92421 DNA enco
9	2759.4	88.4	2797	12 ADL13374	Adl13374 Human ste
10	2742.4	87.9	2929	10 ACA92425	Ac92425 DNA enco
11	2647.4	84.9	2649	14 ADV43981	Adv43981 Human psy
12	2644.4	84.8	2671	6 ABK83322	Abk83322 cDNA enco
13	2644.2	84.7	2649	8 ABK12255	Abk12255 cDNA enco
14	2639.8	84.6	2643	6 AAH99935	Aah99935 Coding se
15	2500.4	80.1	4523	6 ABK83325	Abk83325 cDNA enco
16	2490	79.8	2842	6 ABN59774	Abn59774 Novel hum
17	2315.8	74.2	2696	12 ADI16386	Adi16386 Human pro
18	2208	70.8	2510	6 AAD23843	Aad23843 Human pro
19	2167.4	69.5	2702	13 ADT04072	Adt04072 Human pro

20	2132	68.3	2668	6 ABN59775	Abn59775 Novel hum
21	2056.4	65.9	4309	6 ABK83328	Abk83328 cDNA enco
22	2037.4	65.3	2349	14 ADV43982	Adv43982 Human psy
23	1994.8	63.9	2161	4 AAH15009	Aah15009 Human cDN
c	24	1994	2083	7 ADR41222	Adr41222 Human CD-
25	1827	58.6	2463	10 ACA92424	Ac92424 DNA enco
26	1819.4	58.3	1821	6 ABV76411	Abv76411 Dipeptidy
27	1774.8	56.9	2251	10 ADE79035	Ad79035 Human pro
28	1347	43.2	1669	4 AAC85696	Aac85696 Nucleotid
29	1036	33.2	1197	4 AAC85695	Aac85695 Nucleotid
30	926.4	29.7	1356	6 ABK83326	Abk83326 cDNA enco
31	883.8	28.3	3287	6 AAD38955	Aad38955 Alternati
32	862.6	27.6	2660	6 ADI16689	Adi16689 Human NOV
33	862.6	27.6	2660	6 ADI16687	Adi16687 Human NOV
34	861	27.6	2617	6 ABK83323	Abk83323 cDNA enco
35	861	27.6	2660	12 ADN42341	Adn42341 Human cDN
36	861	27.6	3716	6 ABQ75955	Abq75955 Human PMM
37	861	27.6	4147	12 ADQ86753	Adq86753 Human tum
38	861	27.6	4147	13 ADQ84273	Adq84273 Human tum
39	861	27.6	4147	13 ACN40423	Acn40423 Tumour-as
40	861	27.6	4219	6 ABK83335	Abk83335 cDNA enco
41	861	27.6	4302	6 ABK83333	Abk83333 cDNA enco
42	861	27.6	4398	13 ADS10267	Adsi0267 Human the
43	857.8	27.5	3024	6 AAD38954	Aad38954 Human dip
44	854	27.4	2495	6 AAD38957	Aad38957 Human dip
45	840	26.9	2751	6 AAD38311	Aad38311 Murine di

ALIGNMENTS

RESULT 1
AAC85694
ID AAC85694 standard; cDNA; 3120 BP.

XX AAC85694;

XX 29-JUN-2001 (first entry)

DE Nucleotide sequence of human DPP8.

XX Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;
XX dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
XX growth hormone deficiency; glucose level; mucosal regeneration;
XX non-insulin dependent diabetes mellitus; glucose intolerance;
XX immunosuppression, ss.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 214..2862
XX FT /*tag= a
XX FT /product= "Human DPP8"

XX WO200119866-A1.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-AU001085.

XX 10-SEP-1999; 99AU-00002762.

XX 18-FEB-2000; 2000AU-00005709.

XX (UNSY) UNIV SYDNEY.

XX Abbott CA, Gorell MD;

XX WPI; 2001-281520/29.

XX P-PSDB; AAB47187.

XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving

XX substrates, identifying inhibitors of DPP8 catalytic activity which have

XX therapeutic uses, and for detecting activated T cells.


```
Db 1801 TCCCTTTAGACATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTACACATTCCTGCTGCATCAGTCAGCAGCTGTGACTCTTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTACACATTCCTGCTGCATCAGTCAGCAGCTGTGACTCTTTTATA 1920
Qy 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCTCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCTCT 1980
Qy 1981 GAAGATGACCCAACTTCGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCCAACTTCGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
Qy 2041 CCTCTTCTGCTACTACTCTCCAGAAATTTTCTCTTTTGAAAGTACTACTGGAATTTACA 2100
Db 2041 CCTCTTCTGCTACTACTCTCCAGAAATTTTCTCTTTTGAAAGTACTACTGGAATTTACA 2100
Qy 2101 TTGTAATGGATGCTCTACAAGCCTCATGATCTACAGCTGGGAAAGAAATATCCTACTGTG 2160
Db 2101 TTGTAATGGATGCTCTACAAGCCTCATGATCTACAGCTGGGAAAGAAATATCCTACTGTG 2160
Qy 2161 CTGTTTCATATATGTGTGCTCCTCAGTGCAGTGGTGAATAATCGGTTTAAAGAGTCAAG 2220
Db 2161 CTGTTTCATATATGTGTGCTCCTCAGTGCAGTGGTGAATAATCGGTTTAAAGAGTCAAG 2220
Qy 2221 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGTGTCTAGTGTATAGACACAGG 2280
Db 2221 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGTGTCTAGTGTATAGACACAGG 2280
Qy 2281 GGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
Db 2281 GGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
Qy 2341 GAAATTTGAGATCAGTGGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2341 GAAATTTGAGATCAGTGGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Qy 2401 TTAGATCTGTGGGATCCACGGCTGTCTCTATGAGGATACCTCTCCCTGTATGGCATTA 2460
Db 2401 TTAGATCTGTGGGATCCACGGCTGTCTCTATGAGGATACCTCTCCCTGTATGGCATTA 2460
Qy 2461 ATGCAGAGGTGAGATATCTTCAGGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATC 2520
Db 2461 ATGCAGAGGTGAGATATCTTCAGGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATC 2520
Qy 2521 TTCTATGATACAGGATACACGGAACTTTATATGAGGTCACTCCCTGACCAAGATGAACGGGC 2580
Db 2521 TTCTATGATACAGGATACACGGAACTTTATATGAGGTCACTCCCTGACCAAGATGAACGGGC 2580
Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Db 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Qy 2641 CTGCTCTTACATGTTTCTCTGGATGAGAAATGCAATTTTGACATACCAATGATTTACTG 2700
Db 2641 CTGCTCTTACATGTTTCTCTGGATGAGAAATGCAATTTTGACATACCAATGATTTACTG 2700
Qy 2701 AGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAC 2760
Db 2701 AGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAC 2760
Qy 2761 AGCATAAGGTTCTGATCGGAGAACATTTATGAACTGCTCTTTTGCACTACCTTCAA 2820
Db 2761 AGCATAAGGTTCTGATCGGAGAACATTTATGAACTGCTCTTTTGCACTACCTTCAA 2820
Qy 2821 GAAACCTTTGGATCAGTATTTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
Db 2821 GAAACCTTTGGATCAGTATTTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
Qy 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGA 2940
Db 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGA 2940
```

```
Qy 2941 ATTGATCATCATTTTGATACCTGCGCATGTAAACATCTACTCTGAAAAATAAATGTGGT 3000
Db 2941 ATTGATCATCATTTTGATACCTGCGCATGTAAACATCTACTCTGAAAAATAAATGTGGT 3000
Qy 3001 CCATCGAGGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAA 3060
Db 3001 CCATCGAGGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAA 3060
Qy 3061 TCAATGTATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAAATAAAA 3120
Db 3061 TCAATGTATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAAATAAAA 3120
```

RESULT 2

AAD38956

ID AAD38956 standard; cDNA; 3120 BP.

XX AAD38956;

XX DT 23-SBP-2002 (first entry)

XX DE Human dipeptidyl peptidase 8 (DPP8) cDNA.

XX KW Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
XX KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
XX KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
XX KW antiviral; enzyme; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 214..2862

XX FT /*tag= a

XX FT /product= "Human DPP8 protein"

XX PN WO200234900-A1.

XX XX 02-MAY-2002.

XX XX 29-OCT-2001; 2001WO-AU001388.

XX XX 27-OCT-2000; 2000AU-00001078.

XX PR (UNSY) UNIV SYDNEY.

XX PI Abbott CA, Gorrell MD;

XX XX WPI; 2002-454646/48.

XX DR P-PSDB; AAE24170.

XX XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
XX PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
XX FT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
XX FT infection.

XX PS Example; Fig 1; 91pp; English.

XX CC The present invention relates to dipeptidyl peptidase (DPP) proteins and
XX CC polynucleotides encoding such proteins. The DPP peptides are useful for
XX CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
XX CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
XX CC rejection and HIV (human immuno deficiency virus) infection. The present
XX CC sequence is human DPP8 cDNA

XX SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;

Query Match: 100.0%; Score 3120; DB 6; Length 3120;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGTGTAAAGCTCCGAGGCCGCTGCTACTGCCGCGCTGCTTCTTAGTCCG 60

Db 1 AAGTCCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG 60
Qy 61 GCTTCGCGCGCTGCGGTGTGTCACGCGCGCGCGCGCGAGGAAGCACTGCAACACGAGACGG 120
Db 61 GCTTCGCGCGCTGCGGTGTGTCACGCGCGCGCGCGCGAGGAAGCACTGCAACACGAGACGG 120
Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGAGCGCCGCTCCATAGCGCACTGTCGGGACGG 180
Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGAGCGCCGCTCCATAGCGCACTGTCGGGACGG 180
Qy 181 TCCGCGCGCGCGCGCGAGGAAGGAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
Db 181 TCCGCGCGCGCGCGCGAGGAAGGAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
Qy 241 CTGGGTGTTGAGATATTGAAACTGCGGACTGTGAGGAGAAATGGAATCACAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTGAAACTGCGGACTGTGAGGAGAAATGGAATCACAGGATCGG 300
Qy 301 CCTAAATGGAGCGCTTTTATGTTGAGCGGTATTCTGGAGTCAGCTTAAAGCTGCTT 360
Db 301 CCTAAATGGAGCGCTTTTATGTTGAGCGGTATTCTGGAGTCAGCTTAAAGCTGCTT 360
Qy 361 GCCGATACAGAAATATCATGCTACATGATGGCTAAGGCCACCATGATTTGATGCTTT 420
Db 361 GCCGATACAGAAATATCATGCTACATGATGGCTAAGGCCACCATGATTTGATGCTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGGAACCTCATTTACAGACAGAAATCTATTACCTTGCCATGTCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGAACCTCATTTACAGACAGAAATCTATTACCTTGCCATGTCT 480
Qy 481 GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAAACTATCAATAGAGCA 540
Db 481 GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAAACTATCAATAGAGCA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGAGTCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGAGTCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTATTTCTCGAGAAGAGAACTATTAAAGAGAAAGAAACGCAATGGAACAGTCGGA 660
Db 601 GGAATGTATTTCTCGAGAAGAGAACTATTAAAGAGAAAGAAACGCAATGGAACAGTCGGA 660
Qy 661 ATTTGCTCTTACGATTATCACCAAGGAGTGAACATTTCTGTTCAAGCGGTAGTGA 720
Db 661 ATTTGCTCTTACGATTATCACCAAGGAGTGAACATTTCTGTTCAAGCGGTAGTGA 720
Qy 721 ATTTATCAGTAAGAGTGGAGGCGCAAGGATTTACGCAACGCTTTTAAAGGCCAAT 780
Db 721 ATTTATCAGTAAGAGTGGAGGCGCAAGGATTTACGCAACGCTTTTAAAGGCCAAT 780
Qy 781 CTAGTGGAAAACCTAGTGTGCCAACAATACGGATGGATCCAAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAAAACCTAGTGTGCCAACAATACGGATGGATCCAAAATTTATGCCCGCTGATCCA 840
Qy 841 GACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTTAACATCGTAAACGAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTTAACATCGTAAACGAGAA 900
Qy 901 GAAAGGAGACTCACCTTATGTGCAACAATGAGTACGCAACATGGAAGAGATCCAGATCA 960
Db 901 GAAAGGAGACTCACCTTATGTGCAACAATGAGTACGCAACATGGAAGAGATCCAGATCA 960
Qy 961 GCTGGAGTCGCTACCTTTGTTCTCAAGAGAAATTTGATAGATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCAAGAGAAATTTGATAGATTTCTGGCTATTGGTGG 1020
Qy 1021 TCTCCAAAAGCTGGAACAACTCCAGTGGTGGTAAATTTCTTAAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAAGCTGGAACAACTCCAGTGGTGGTAAATTTCTTAAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGATCTGAGGTGGAATTTATCATGTTTACATCCCTATGTTGGAACAGAGG 1140
Db 1081 AATGATGATCTGAGGTGGAATTTATCATGTTTACATCCCTATGTTGGAACAGAGG 1140

Qy 1141 GCAGATTCATTTCCGTTATCTCTAAACACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCATTTCCGTTATCTCTAAACACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATG 1200
Qy 1201 TCAGAAAATAATGATTGATGCTTGAAGGAAGGATCATAGATGTCTAGATAAGGAACATAAT 1260
Db 1201 TCAGAAAATAATGATTGATGCTTGAAGGAAGGATCATAGATGTCTAGATAAGGAACATAAT 1260
Qy 1261 CAACCTTTTGAGATCTATTGAAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCT 1320
Db 1261 CAACCTTTTGAGATCTATTGAAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCT 1320
Qy 1321 GAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG 1380
Db 1321 GAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG 1380
Qy 1381 TTGATCTCACTGAAATATTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440
Db 1381 TTGATCTCACTGAAATATTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440
Qy 1441 ATTTGAGTCAGTCCCTGATTTCTGTGACGCCACTTAATTATCTATGAAGAAACAAACAGACATC 1500
Db 1441 ATTTGAGTCAGTCCCTGATTTCTGTGACGCCACTTAATTATCTATGAAGAAACAAACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTTCATGTTTCCCCAAAGTCACGAGAGGAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTCCCCAAAGTCACGAGAGGAATTTGAG 1560
Qy 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
Qy 1621 TTTAAAGGAAGCAAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Db 1621 TTTAAAGGAAGCAAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Qy 1681 TGTCTCTATCAAAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGCCCGCAT 1740
Db 1681 TGTCTCTATCAAAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGCCCGCAT 1740
Qy 1741 GGATCTAATATCCAAAGTTGATGAAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAAGTTGATGAAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCCTGGAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCACTAGTCAGCACTGTCGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCACTAGTCAGCACTGTCGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCT 1980
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCT 1980
Qy 1981 GAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCGCACATTTTGGATTTGACGAGGT 2040
Db 1981 GAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCGCACATTTTGGATTTGACGAGGT 2040
Qy 2041 CCTCTTCTGACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
Db 2041 CCTCTTCTGACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
Qy 2101 TTGATGGGATGCTCTTACAAGCTCATGATCTACAGCTCGGAAAGAAATATCTTACTGTG 2160
Db 2101 TTGATGGGATGCTCTTACAAGCTCATGATCTACAGCTCGGAAAGAAATATCTTACTGTG 2160
Qy 2161 CTGTTTCATATATGTTGGTCTCCTCAGGTGAGTTGGTGAATATCGGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTTCATATATGTTGGTCTCCTCAGGTGAGTTGGTGAATATCGGTTTAAAGGAGTCAAG 2220

QY 2221 TATTTCCGCTTGAATACCTAGAGCTCTTAGGTTATGTGGTTGTAGTATAGAACACAGG 2280
Db 2221 TATTTCCGCTTGAATACCTAGAGCTCTTAGGTTATGTGGTTGTAGTATAGAACACAGG 2280
QY 2281 GGATCTCTGACCCAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATA 2340
Db 2281 GGATCTCTGACCCAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATA 2340
QY 2341 GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2341 GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
QY 2401 TTAGATCGTGTGGGCATCCAGCGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
Db 2401 TTAGATCGTGTGGGCATCCAGCGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
QY 2461 ATGCAAGGTGAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATC 2520
Db 2461 ATGCAAGGTGAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATC 2520
QY 2521 TTCTATGATACAGGATACAGGAACGTTATATGGGTCACTGACAGAGATGAACAGGGC 2580
Db 2521 TTCTATGATACAGGATACAGGAACGTTATATGGGTCACTGACAGAGATGAACAGGGC 2580
QY 2581 TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Db 2581 TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
QY 2641 CTGCTCTTACATGGTTCTCGATGAGATGTCCATTTGACATACAGATATATCTG 2700
Db 2641 CTGCTCTTACATGGTTCTCGATGAGATGTCCATTTGACATACAGATATATCTG 2700
QY 2701 AGTTTCTTGTAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACAC 2760
Db 2701 AGTTTCTTGTAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACAC 2760
QY 2761 AGCATAAGAGTTCTCGAATCGGAGAACATTTATGAACTGATCTTTTGCACTACCTTCAA 2820
Db 2761 AGCATAAGAGTTCTCGAATCGGAGAACATTTATGAACTGATCTTTTGCACTACCTTCAA 2820
QY 2821 GAAACCTTTGGATCAGTATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
Db 2821 GAAACCTTTGGATCAGTATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
QY 2881 TCTCTGGTATACATCTGGCTATTTAACCAATGAGGAGTTTATCAACAGAAACACAGA 2940
Db 2881 TCTCTGGTATACATCTGGCTATTTAACCAATGAGGAGTTTATCAACAGAAACACAGA 2940
QY 2941 ATTGATCATCATTTTGTATACCTGCGCATGTAACTCTCTCTGAAATAAATGTGGTG 3000
Db 2941 ATTGATCATCATTTTGTATACCTGCGCATGTAACTCTCTCTGAAATAAATGTGGTG 3000
QY 3001 CCATGCGAGGGTCTACGGTTTGTGTAGTAACTTAATACCTTAAACCCCATGCTCAAAA 3060
Db 3001 CCATGCGAGGGTCTACGGTTTGTGTAGTAACTTAATACCTTAAACCCCATGCTCAAAA 3060
QY 3061 TCAATATGATACATATCTGTAGAGAGCCAGCAATACCATAGAAATTAATAAATAAATAA 3120
Db 3061 TCAATATGATACATATCTGTAGAGAGCCAGCAATACCATAGAAATTAATAAATAAATAA 3120

RESULT 3
ABK83327
ID ID
XX ABK83327 standard; cDNA; 4829 BP.
XX
AC ABK83327;
XX
XX
DT 12-AUG-2002 (first entry)
XX
DE cDNA encoding human DPRP-1 splice variant #3.
XX
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW

KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
XX metabolic disorder; gene; ss.
OS Homo sapiens.
XX
PN WO200231134-A2.
XX
XX 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031874.
XX
XX 12-OCT-2000; 2000US-0240117P.
PR
XX (FERR) FERRING BV.
PA
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
PI P-PSDB; ABG61596.
DR
DR WPI; 2002-444178/47.
XX
XX
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
XX Disclosure; Page 65-66; 113pp; English.
PS
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraines, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX
SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;
Query Match 99.6%; Score 3106.4; DB 6; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1 AAGTGCTAAAGCTCCGAGGCCAAGCGCGCTACTGCGCGCGCTGCTTCTTAGTCCG 60
Db 1 AAGTGCTAAAGCTCCGAGGCCAAGCGCGCTACTGCGCGCGCTGCTTCTTAGTCCG 60
QY 61 CGTTCCGCGCTGGTGTTCACCGCGCGCGCGCGAGAGGCCACTGCAACACGAGCCG 120
Db 61 CGTTCCGCGCTGGTGTTCACCGCGCGCGCGCGAGAGGCCACTGCAACACGAGCCG 120
QY 121 GAGTGAGGCGCGCGAGCATGAGCGCGCGCGCGCGCTCCATAGCGACCTCGGCGCG 180
Db 121 GAGTGAGGCGCGCGAGCATGAGCGCGCGCGCGCGCTCCATAGCGACCTCGGCGCG 180
QY 181 TCCGGGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACACAGAAC 240
Db 181 TCCGGGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACACAGAAC 240
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAAGATCG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAAGATCG 300
QY 301 CCTAAATGAGCGCTTTTATGTTAGCGGTATTCTCGAGTCAGCTTAAAGAGCTGCTT 360
Db 301 CCTAAATGAGCGCTTTTATGTTAGCGGTATTCTCGAGTCAGCTTAAAGAGCTGCTT 360

Qy	361	GCCGATACAGAAATATCATGCTATGCGCTAAGGCACCAATGATTTTCATGTTT	420	1441	ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATATTCTATGAAGAAACAAACAGACATC	1500
Db	361			1441		
Qy	421	GTGAAGAGAAATGATCCAGATGGAACCTCATTTACAGACGAATCTATTACCTTCCCATGCT	480	1501	TGGATTAATATCCATGACATCTTTTCATGTTTTTCCCAAAGTCACGAGAGGAATTTGAG	1560
Db	421			1501		
Qy	481	GTTGAGAACAGAGAAATACACTGTTTTATTCTGAAATTTCCCAAACTATCAATAGAGCA	540	1561	TTTTTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT	1620
Db	481			1561		
Qy	541	GCAGTCTTAATGCTCTCTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600	1621	TTAAAGGAAAGCAAAATATAAAACGATCCAGTGGTGGCTGCTGCTCAAGTGATTTCAAG	1680
Db	541			1621		
Qy	601	GGAATGTATTCTCGAGAAGAGAACTATTAAAGAGAAAGAAACGCAATTTGAAACAGTCGGA	660	1681	TGTCTCTATCAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	601			1681		
Qy	661	ATTGCTCTTACGATTATACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA	720	1741	GGATCTTAATATCCAAAGTTGATCAAGGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Db	661			1741		
Qy	721	ATTTATACGTAAGATGGAAGGCCACRAAGGATTTACGCAACACTTTTAAGGCCCAAT	780	1801	TCCCTTTTAGAGCATCACCTGCTAGCTAGTACGTAAATCCTGAGAGGTCGACAAAGG	1860
Db	721			1801		
Qy	781	CTAGTGGAACTAGTGTGTCACATACGATCGATCCAAATTTATGCCCGCTGATCCA	840	1861	CTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGCACTGTGACTTCTTTATA	1920
Db	781			1861		
Qy	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTGTAACGAGAA	900	1921	AGTAAATAGTAAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCCCT	1980
Db	841			1921		
Qy	901	GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960	1981	GAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCATTTTGGATTCAGCAGGT	2040
Db	901			1981		
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020	2041	CCTTTCTGACATACTACTCTCCAGAAATTTTCTTTTGAAGAGTACTACTGGAATTACA	2100
Db	961			2041		
Qy	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA	1080	2101	TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCTCGGAAAGAAATATCCTACTGTG	2160
Db	1021			2101		
Qy	1081	AATGATGAATCTGAGGTGGAATTTATTCTAGTTACATCCCTCATGTTGGAACAAGGAGG	1140	2161	CTGTTTCATATATGTTGG--TCTCTCAGGTGCAAGTTGGTGAATTAATTCGGTTTAAAGGAGTCA	2218
Db	1081			2161		
Qy	1141	GCAGATTCATTCGGTTATCTTAACAGGTACAGCAATCCTAAAGTCACTTTTAAGATG	1200	2219	AGTATTTCCGCTTGAAATACCCCTAGCTCTCTAGGTATATGTGTTGTAGTGATAGACAACA	2278
Db	1141			2219		
Qy	1201	TCAGAAATATGATGATGCTGGAAGAGATCATAGATGCTATAGATGAAGAACTAATTT	1260	2281	GGGGATCCTGTCAACAGGGCTTAAATTTGAAGCGCTTTTAAATATAAAAATGGGTCAAA	2338
Db	1201			2281		
Qy	1261	CAACCTTTTGAGATTTCTTTTGAAGAGTTGGAATATATTCGCGAGCTGGATGGACTCCT	1320	2339	TAGAAATGACCATCAGGTGGAGGACCTCAATATCTAGCTTCTCGCATATGATTTTCAATG	2398
Db	1261			2339		
Qy	1321	GAGGAAATATGCTTGGTCCATCTATAGATCGCTCCAGACTCGCCTACAGATAGTG	1380	2341	TAGAAATGACCATCAGGTGGAGGACCTCAATATCTAGCTTCTCGCATATGATTTTCAATG	2400
Db	1321			2341		
Qy	1381	TTGATCTCACCTGAAATTTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC	1440	2399	ACTTAGTCTGTGGGCATCCAGCGTGTCTATGAGGATACCTCTCCTCATGATGCGCAT	2458
Db	1381			2399		
Qy	1441	TTGATCTCACCTGAAATTTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC	1500	2401	ACTTAGTCTGTGGGCATCCAGCGTGTCTATGAGGATACCTCTCCTCATGATGCGCAT	2460
Db	1441			2401		
Qy	1501	TTGATCTCACCTGAAATTTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC	1560	2459	TAATGAGAGGTCAGATATCTTCAGGGTTGCTATTTGCTGGGCGCCAGTCCTCTGTGGA	2518
Db	1501			2459		
Qy	1561	TTGATCTCACCTGAAATTTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC	1620	2461	TAATGAGAGGTCAGATATCTTCAGGGTTGCTATTTGCTGGGCGCCAGTCCTCTGTGGA	2520
Db	1561			2461		
Qy	1621	TTGATCTCACCTGAAATTTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC	1680	2519	TCTTCTATGATACAGGATACAGGAAACGTTTATATGGGTCAACCTGACCAAGATGAACAGG	2578
Db	1621			2519		

Db	2521	TCTTCTATGATACAGGATACAGGAACGGTTATATGGTCCACCTGACCAAGTAAACAGG	2580
Qy	2579	GCTATTAAGTATGAGTCTGGCCATGCAAGCAGAAAGTCCCTCTGAAACCAATCGTT	2638
Db	2581	GCTATTAAGTATGAGTCTGGCCATGCAAGCAGAAAGTCCCTCTGAAACCAATCGTT	2640
Qy	2639	TACTGCTCTTACATGTTTCTCGATGAGAAATGCTCAATTTTGACATACCAAGTATATTAC	2698
Db	2641	TACTGCTCTTACATGTTTCTCGATGAGAAATGCTCAATTTTGACATACCAAGTATATTAC	2700
Qy	2699	TGAGTCTTTTGTAGTGAGGCTGGAAAGCCATATGATTACAGATCTATCTCTAGAGAGAC	2758
Db	2701	TGAGTCTTTTGTAGTGAGGCTGGAAAGCCATATGATTACAGATCTATCTCTAGAGAGAC	2760
Qy	2759	ACAGATAAGAGTTCTGAAATCGGGAGAACATATGAACTGATCTTTTGCACTACCTTC	2818
Db	2761	ACAGATAAGAGTTCTGAAATCGGGAGAACATATGAACTGATCTTTTGCACTACCTTC	2820
Qy	2819	AAGAAACCTTGGATCAGTATGCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGA	2878
Db	2821	AAGAAACCTTGGATCAGTATGCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGA	2880
Qy	2879	ACTCTCTGTATACACTGGCTATTATTAACCAATGAGGAGTTTAAATCAACAGAAACACA	2938
Db	2881	ACTCTCTGTATACACTGGCTATTATTAACCAATGAGGAGTTTAAATCAACAGAAACACA	2940
Qy	2939	GAAATGATCATCACATTTTGTATCTGCTGCAATCAATCTCTGCTGAAATTAATGTGG	2998
Db	2941	GAAATGATCATCACATTTTGTATCTGCTGCAATCAATCTCTGCTGAAATTAATGTGG	3000
Qy	2999	TGCCATGAGGGCTAGGTTTGTGTAGTATCTTAATCACTTAACCTTAAACCCACATGCTAA	3058
Db	3001	TGCCATGAGGGCTAGGTTTGTGTAGTATCTTAATCACTTAAACCTTAAACCCACATGCTAA	3060
Qy	3059	RATCAAAATGATACATATTTCTTGAGAGACCAGCAATACCAATTAAGTAATTAATAAAAAA	3118
Db	3061	RATCAAAATGATACATATTTCTTGAGAGACCAGCAATACCAATTAAGTAATTAATAAAAAA	3120
Qy	3119	AA 3120	
Db	3121	AA 3122	
RESULT 4			
ID	AAH99934	standard; cDNA; 3143 BP.	
XX	AAH99934;		
AC	AAH99934;		
DT	12-APR-2002	(first entry)	
XX	cDNA encoding 21953 human prollyl oligopeptidase.		
XX	21953 prollyl oligopeptidase; human; proline; endopeptidase; cancer;		
KW	cardiovascular disease; autoimmune disease; atopic allergy;		
KW	neural disorder; vascular disorder; prostate disorder; cytostatic;		
KW	antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;		
KW	diabetes mellitus; arthritis; multiple sclerosis; asthma;		
KW	Grave's disease; neuronal disorder; demyelinating disease; ss.		
XX	Homo sapiens.		
OS			
XX			
XX			
PH	Key	Location/Qualifiers	
FT	CDS	229..2877	
FT		/tag= a	
FT		/product= "21953 prollyl oligopeptidase"	
FT		/note= "This region is specifically claimed in claim 2"	
XX			
XX	WO200179473-A2.		
XX			
PD	25-OCT-2001.		
XX			

PF	11-APR-2001;	2001WO-US040483.	
XX	PR	18-APR-2000;	2000US-0197508P.
XX	PA	(MILL-) MILLENNIUM PHARM INC.	
XX	PI	Meyers RA, Williamson M;	
XX	XX	WPI; 2002-034353/04.	
DR	DR	P-PSDB; AAG78415.	
XX	XX	New polypeptides 21953, member of human prollyl oligopeptidase family,	
PT	PT	useful as diagnostic targets and therapeutic agents for controlling	
PT	PT	cancer, lymphoma and leukemia.	
XX	XX	Claim 7; Page 100-102; 121pp; English.	
PS	PS	This invention relates to an isolated 21953 human prollyl oligopeptidase.	
CC	CC	Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,	
CC	CC	antithyroid, dermatological, antipsoriatic, antiasthmatic,	
CC	CC	ophthalmological, antiinflammatory, vasotropic, antianginal, cardiant,	
CC	CC	anticonvulsant, gynaecological, anorectic and metabolic in its action. Uses include	
CC	CC	antitherosclerotic, gene therapy, expression or activity of 21953 protein modulator, it is	
CC	CC	useful for identifying a compound which binds to it and can be used in	
CC	CC	preventing, treating or detecting a cellular proliferative or	
CC	CC	differentiative disorder. The 21953 molecules can act as novel diagnostic	
CC	CC	targets and therapeutic agents for controlling disorders associated with	
CC	CC	the aberrant activity or degradation of peptide hormones e.g., disorders	
CC	CC	associated with cell differentiation and proliferation such as cancer,	
CC	CC	immune function, reproductive, neurological and cardiovascular function.	
CC	CC	The 21953 molecules are thus useful for treating and preventing cellular	
CC	CC	proliferative and differentiative disorders, haematopoietic neoplastic	
CC	CC	disorders, immune disorders such as autoimmune diseases, diabetes	
CC	CC	mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,	
CC	CC	neural disorders, demyelinating diseases, vascular disorders and	
CC	CC	metabolism or pain disorders. This sequence represents the cDNA encoding	
CC	CC	sequence of 21953 human prollyl oligopeptidase	
XX	XX	Sequence 3143 BP; 943 A; 644 C; 712 G; 844 T; 0 U; 0 Other;	
SQ	SQ	Query Match 99.5%; Score 3105.2; DB 6; Length 3143;	
		Best Local Similarity 99.9%; Pred. No. 0;	
		Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Qy	11	GCCTCCGAGGCAAGGCGCTGCTACTGCGCGCGCTTCTTAGTGCCTGCTTCCGCGG 70	
Db	26	GCCTCCGAGGCAAGGCGCTGCTACTGCGCGCGCTTCTTAGTGCCTGCTTCCGCGG 85	
Qy	71	CTGGGTTGTACCG 130	
Db	86	CTGGGTTGTACCG 145	
Qy	131	GGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 190	
Db	146	GGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 205	
Qy	191	GCCTGGGGGAGGAGAAATCAACNTGGCAGCAGCAATGGAACAGACAGCTGGTGTG 250	
Db	206	GCCTGGGGGAGGAGAAATCAACNTGGCAGCAGCAATGGAACAGACAGCTGGTGTG 265	
Qy	251	AGATATTTGAACTGCGGAGCTGTGAGGAGAAATTTGAATTCAGAGATCGGCTAAAT 310	
Db	266	AGATATTTGAACTGCGGAGCTGTGAGGAGAAATTTGAATTCAGAGATCGGCTAAAT 325	
Qy	311	AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGTCTGTCGCGATACCA 370	
Db	326	AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGTCTGTCGCGATACCA 385	
Qy	371	GAAATATCATGCTACATGCTGCTTAAGCCACCATGATTTTCATGTTTGTGAAGAGA 430	
Db	386	GAAATATCATGCTACATGCTGCTTAAGCCACCATGATTTTCATGTTTGTGAAGAGA 445	

QY 431 ATGATCCAGATGGACCTCATTTCAGACAGAATCTATTACCTTGCCATGTCGTGTGAGAACA 490
Db 446 ATGATCCAGATGGACCTCATTTCAGACAGAATCTATTACCTTGCCATGTCGTGTGAGAACA 505
QY 491 GAGAAATACACTGTTTATTCTGAAATCCCAAACTATCAATAGACGACGCTTAA 550
Db 506 GAGAAATACACTGTTTATTCTGAAATCCCAAACTATCAATAGACGACGCTTAA 565
QY 551 TGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTTGAACACTTGAATGTATT 610
Db 566 TGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTTGAACACTTGAATGTATT 625
QY 611 CTCGAGAAAGAACTATTAAAGAAAGAAACGCAATGGAACACAGTCGGAAATGCTTCTT 670
Db 626 CTCGAGAAAGAACTATTAAAGAAAGAAACGCAATGGAACACAGTCGGAAATGCTTCTT 685
QY 671 ACGATTATCACAAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACAG 730
Db 686 ACGATTATCACAAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACAG 745
QY 731 TAAAGATGGAGGCCCAAGGATTTACGCAACAACTTTAAAGGCCCAATCTAGTGGAAA 790
Db 746 TAAAGATGGAGGCCCAAGGATTTACGCAACAACTTTAAAGGCCCAATCTAGTGGAAA 805
QY 791 CTAGTTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTG 850
Db 806 CTAGTTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTG 865
QY 851 CTTTATTATACATAGCAACGATATTGGATATCTAAACATCGTAACAGAGAAAGAGAC 910
Db 866 CTTTATTATACATAGCAACGATATTGGATATCTAAACATCGTAACAGAGAAAGAGAC 925
QY 911 TCACCTATGTGCAATGAGCTAGCCAACTGGAAGAAAGATGCCAGATCAGCTGGAGTCG 970
Db 926 TCACCTATGTGCAATGAGCTAGCCAACTGGAAGAAAGATGCCAGATCAGCTGGAGTCG 985
QY 971 CTACCTTTGTTCTCAAGAGAAATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAG 1030
Db 986 CTACCTTTGTTCTCAAGAGAAATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAG 1045
QY 1031 CTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGAAT 1090
Db 1046 CTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGAAT 1105
QY 1091 CTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGAAAACAAAGGAGGCGAGATTCAAT 1150
Db 1106 CTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGAAAACAAAGGAGGCGAGATTCAAT 1165
QY 1151 TCCGTTATCCCTAAACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAA 1210
Db 1166 TCCGTTATCCCTAAACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAA 1225
QY 1211 TGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACTAAATTCACACCTTTTG 1270
Db 1226 TGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACTAAATTCACACCTTTTG 1285
QY 1271 AGATTCTATTGGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAAT 1330
Db 1286 AGATTCTATTGGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAAT 1345
QY 1331 ATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC 1390
Db 1346 ATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC 1405
QY 1391 CTGAATATTATTATCCAGTGAAGATGATTATGGAAGGCGAGAGACTATTGATGTCAG 1450
Db 1406 CTGAATATTATTATCCAGTGAAGATGATTATGGAAGGCGAGAGACTATTGATGTCAG 1465
QY 1451 TGCCGTGATCTGTGACGCCCACTAAATTTATCTATGAAGAAACAAACAGACATCTGGATAATA 1510
Db 1466 TGCCGTGATCTGTGACGCCCACTAAATTTATCTATGAAGAAACAAACAGACATCTGGATAATA 1525
QY 1511 TCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAGAAATTCAGTTTATTTTGTG 1570

Db 1526 TCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAGAAATTCAGTTTATTTTG 1585
QY 1571 CCTCTGAATGCAAAACAGGTTTCGTCTATTTATACAAATTTACATCTATTTTAAAGGAAA 1630
Db 1586 CCTCTGAATGCAAAACAGGTTTCGTCTATTTATACAAATTTACATCTATTTTAAAGGAAA 1645
QY 1631 GCAATATATAAGCATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCA 1690
Db 1646 GCAATATATAAGCATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCA 1705
QY 1691 AAGAGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATA 1750
Db 1706 AAGAGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATA 1765
QY 1751 TCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACAAAGACTCCCTTTTAG 1810
Db 1766 TCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACAAAGACTCCCTTTTAG 1825
QY 1811 AGCATCACCTGTATCGTAGTCAAGTTCAGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACC 1870
Db 1826 AGCATCACCTGTATCGTAGTCAAGTTCAGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACC 1885
QY 1871 GTGGCTACTCAATTTCTGTCATCAGTCAAGTCTGCTGCTCTTTTATTAAGTAAGTATA 1930
Db 1886 GTGGCTACTCAATTTCTGTCATCAGTCAAGTCTGCTGCTCTTTTATTAAGTAAGTATA 1945
QY 1931 GTAAACAGAGAATCCACACTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACC 1990
Db 1946 GTAAACAGAGAATCCACACTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACC 2005
QY 1991 CAACTTGCANAAACAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTCTCCCTG 2050
Db 2006 CAACTTGCANAAACAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTCTCCCTG 2065
QY 2051 ACTATACTCTCCAGAAATTTCTTTTGAAGTACTACTGGATTTACATTTGATGGA 2110
Db 2066 ACTATACTCTCCAGAAATTTCTTTTGAAGTACTACTGGATTTACATTTGATGGA 2125
QY 2111 TGCTCTACAGCTCATGATCTACAGCTGGAAAGAAATTCCTACTGTGCTGTTTCATAT 2170
Db 2126 TGCTCTACAGCTCATGATCTACAGCTGGAAAGAAATTCCTACTGTGCTGTTTCATAT 2185
QY 2171 ATGGTGGTCTCAGGTGCAAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTCGGCT 2230
Db 2186 ATGGTGGTCTCAGGTGCAAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTCGGCT 2245
QY 2231 TGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAAGGGGATCCTGTC 2290
Db 2246 TGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAAGGGGATCCTGTC 2305
QY 2291 ACCGAGGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAAATAGAAATTTGACG 2350
Db 2306 ACCGAGGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAAATAGAAATTTGACG 2365
QY 2351 ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTG 2410
Db 2366 ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTG 2425
QY 2411 TGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGGATTAATCCAGAGT 2470
Db 2426 TGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGGATTAATCCAGAGT 2485
QY 2471 CAGATATCTTCAGGGTTGCTATTCCTGGGCGCCAGTCACTCTGTGGATCTCTATGATA 2530
Db 2486 CAGATATCTTCAGGGTTGCTATTCCTGGGCGCCAGTCACTCTGTGGATCTCTATGATA 2545
QY 2531 CAGGATACCGGAAGTATTATGGGTCACTGTGACAGATGAACAGGGCTATTACTTAG 2590
Db 2546 CAGGATACCGGAAGTATTATGGGTCACTGTGACAGATGAACAGGGCTATTACTTAG 2605
QY 2591 GATCTGTGGCATGCAAGCAAGAAAGTTCCCTCTCTGAAACCAATTCGTTTACTGCTCTTAC 2650

Db 2606 GATCTGTGGCCATGCAAGCAGAAAGATTTCCCTCTCGAACCAAAATCGTTTACTGCTCTTAC 2665
Qy 2651 ATGGTTTCTCTGATGAGAATGTCATTTTGCACATACCAAGTATATTTACTGAGTCTTTTATG 2710
Db 2666 ATGGTTTCTCTGATGAGATGTCATTTTGCACATACCAAGTATATTTACTGAGTCTTTTATG 2725
Qy 2711 TGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGACACAGCATTAAGAG 2770
Db 2726 TGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGACACAGCATTAAGAG 2785
Qy 2771 TTCCTGAATCGGAGACATATGACATGCTCTTTTGGCACTACCTTTCAAGAAAACCTTTG 2830
Db 2786 TTCCTGAATCGGAGACATATGACATGCTCTTTTGGCACTACCTTTCAAGAAAACCTTTG 2845
Qy 2831 GATCAGGTATTTGCTCTTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGGTAT 2890
Db 2846 GATCAGGTATTTGCTCTTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGGTAT 2905
Qy 2891 ACATGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATCATC 2950
Db 2906 ACATGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATCATC 2965
Qy 2951 ACATTTGATACCTGCCATGAAATCACTACTCTGAAATTAATGTTGGTCCATGCAGGG 3010
Db 2966 ACATTTGATACCTGCCATGAAATCACTACTCTGAAATTAATGTTGGTCCATGCAGGG 3025
Qy 3011 GTCTACGGTTTGTGTAGTATCTAATTAACCTTAAACCCACATCTCAAAATCAAAATGATA 3070
Db 3026 GTCTACGGTTTGTGTAGTATCTAATTAACCTTAAACCCACATCTCAAAATCAAAATGATA 3085
Qy 3071 CATATTTCTGAGAGACCCAGCAATACCAATGAATTAATAAAAAA 3120
Db 3086 CATATTTCTGAGAGACCCAGCAATACCAATGAATTAATAAAAAA 3135

RESULT 5
ABK12892
ID ABK12892 standard; cDNA; 3106 BP.
XX
AC ABK12892;
XX
DT (first entry)
XX
DE Human protease PRTS-9 cDNA sequence.
XX
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 203..2851
FT /*tag= a
FT /product= "Human protease PRTS-9"
XX
WO200198468-A2.
PN
XX 27-DEC-2001.
PD
XX 13-JUN-2001; 2001WO-US019178.
XX
PF 16-JUN-2000; 2000US-0212336P.
XX
PR 22-JUN-2000; 2000US-0213955P.
PR
PR 29-JUN-2000; 2000US-0215396P.
PR
PR 07-JUL-2000; 2000US-0216821P.
PR
PR 14-JUL-2000; 2000US-0218946P.
XX
PA (INCYTE) INCYTE GENOMICS INC.
XX

PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
PI Deleane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
PI Walia NK, Yao MG, Lu DM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
XX
DR WPI; 2002-090437/12.
DR P-PSDB; AAU74749.
XX
PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
PT (e.g. cancer) disorders.
XX
PS Claim 5; Page 166-167; 177pp; English.
XX
CC The present invention relates to twenty one new human proteases, referred
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.
CC The present nucleic acid sequence encodes the human protease PRTS-9
CC protein of the invention
XX
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;
Query Match 99.4%; Score 3101.8; DB 6; Length 3106;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 13 CTCGAGGCCAAGGCGCTGTCTACTGCGCGCGCTTCTTAGTCCGCGTTCGCGCGCT 72
Db 2 CTCGAGGCCAAGGCGCTGTCTACTGCGCGCGCTTCTTAGTCCGCGTTCGCGCGCT 61
Qy 73 GGGTTCTACCG 132
Db 62 GGGTTCTACCG 121
Qy 133 CGCAGCATGAAGCG 192
Db 122 CGCAGCATGAAGCG 181
Qy 193 CGGGGGGAAGGAAATGCAATGCGAGCGAGCAATGCGAAACAGAACAGAGCTGGTGTGAG 252
Db 182 CGGGGGGAAGGAAATGCAATGCGAGCGAGCAATGCGAAACAGAACAGAGCTGGTGTGAG 241
Qy 253 ATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAGAGGATCGGCTAAATTTGAG 312
Db 242 ATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAGAGGATCGGCTAAATTTGAG 301
Qy 313 CTTTTTATGTTGAGCGGATTTCTTGGAGTCAAGCTTAAAGAGCTTCCGCGATACCAGA 372
Db 302 CTTTTTATGTTGAGCGGATTTCTTGGAGTCAAGCTTAAAGAGCTTCCGCGATACCAGA 361
Qy 373 AAATATCATGGCTTACATGATGCTAAGGCAACCAATGATTTTGTGAAGGAGAT 432
Db 362 AAATATCATGGCTTACATGATGCTAAGGCAACCAATGATTTTGTGAAGGAGAT 421
Qy 433 GATCCAGATGAGCTCATTCAGACAGATCTATTACCTTGCATGCTGCTGGAGACAGA 492
Db 422 GATCCAGATGAGCTCATTCAGACAGATCTATTACCTTGCATGCTGCTGGAGACAGA 481
Qy 493 GAAATATACCTGTTTATTTCTGAAATTTCCAAAACCTATCAATAGAGAGAGCTTTAATG 552
Db 482 GAAATATACCTGTTTATTTCTGAAATTTCCAAAACCTATCAATAGAGAGAGCTTTAATG 541
Qy 553 CTCTCTTGGAGGCTCTCTTTTGGATCTTTTTCAGGCAACACTGGAATGATTTCT 612

[illegible]

QY 2773 CCTGAATCGGAGAACATTTATGAATGCTATCTTTTGCACCTACCTTCAAGAAACCTTGA 2832
DB 2762 CCTGAATCGGAGAACATTTATGAATGCTATCTTTTGCACCTACCTTCAAGAAACCTTGA 2821
QY 2833 TCACGTATGCTGCTCTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATAC 2892
DB 2822 TCACGTATGCTGCTCTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATAC 2881
QY 2893 ACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATTCATCATCAC 2952
DB 2882 ACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATTCATCATCAC 2941
QY 2953 ATTTTGATACCTGCGCATGTAACATCTACTCTCTGAAATAAATGTGGTGCATCGAGGGT 3012
DB 2942 ATTTTGATACCTGCGCATGTAACATCTACTCTCTGAAATAAATGTGGTGCATCGAGGGT 3001
QY 3013 CTACGGTTTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACA 3072
DB 3002 CTACGGTTTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACA 3061
QY 3073 TATTCCTGAGAGCCAGCAATACCAATGAAGTAATTAACAAAAA 3117
DB 3062 TATTCCTGAGAGCCAGCAATACCAATGAAGTAATTAACAAAAA 3106

RESULT 6

ABK83332
ID ABK83332 standard; cDNA; 4685 BP.

AC ABK83332;

DT 12-AUG-2002 (first entry)

DE cDNA encoding human DPRP-1 splice variant #8.

KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.

XX Homo sapiens.

XX WO200231134-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US031874.

XX 12-OCT-2000; 2000US-0240117P.

XX (FERR) FERRING BV.

XX Qi S, Akinsanya KO, Riviere PJ, Junien J;

XX WPI; 2002-444178/47.

XX P-PSDB; ABG61601.

PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.

XX Disclosure; Page 75-76; 113pp; English.

XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,

CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins

SQ Sequence 4685 BP; 1430 A; 853 C; 991 G; 1411 T; 0 U; 0 Other;

Query Match 90.5%; Score 2824.4; DB 6; Length 4685;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;

QY 1 AAGTGTCTAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCCGCTGCTTCTTAGTGGCG 60
DB 1 AAGTGTCTAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCCGCTGCTTCTTAGTGGCG 60
QY 61 CGTTCGCCGCTGGTGTGTACCGCGCGCGCGCGAGGAAGCCACTGCAACAGGACCG 120
DB 61 CGTTCGCCGCTGGTGTGTACCGCGCGCGCGCGAGGAAGCCACTGCAACAGGACCG 120
QY 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCGCTCGGGACGG 180
DB 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCGCTCGGGACGG 180
QY 181 TCCGGCGCGCGCGGGGGAAGAAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
DB 181 TCCGGCGCGCGCGGGGGAAGAAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTCGGACTGTGAGGAGATATTTGAATCAAGATCGG 300
DB 241 CTGGGTGTTGAGATATTTGAAACTCGGACTGTGAGGAGATATTTGAATCAAGATCGG 300
QY 301 CCTAAATTTGAGGCTTTTATGTTGAGCGGTATTTCTGAGCTCAGCTTAAAAGCTGTT 360
DB 301 CCTAAATTTGAGGCTTTTATGTTGAGCGGTATTTCTGAGCTCAGCTTAAAAGCTGTT 360
QY 361 GCCGATACCAAGAAATATCATGCTACATGATGGCTTAAGGACCAACATGATTTTATGTTT 420
DB 361 GCCGATACCAAGAAATATCATGCTACATGATGGCTTAAGGACCAACATGATTTTATGTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTATTACCTTGCCATGCT 480
DB 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTATTACCTTGCCATGCT 480
QY 481 GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
DB 481 GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
QY 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
DB 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
QY 601 GGAATGTATTCTCGAGAAGAACTATTATTAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
DB 601 GGAATGTATTCTCGAGAAGAACTATTATTAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
QY 661 ATTGCTTTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGCTAGTGGGA 720
DB 661 ATTGCTTTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGCTAGTGGGA 720
QY 721 ATTTATCACGTAAAGATGGAGGCGGCACCAAGGATTTAGCACAACCTTTAAGGCCCAAT 780
DB 721 ATTTATCACGTAAAGATGGAGGCGGCACCAAGGATTTAGCACAACCTTTAAGGCCCAAT 780
QY 781 CTAGTGAACACTAGTGTGTTCCCAACATACGATGGATCCAAATTTATGCTGCTGATCCA 840
DB 781 CTAGTGAACACTAGTGTGTTCCCAACATACGATGGATCCAAATTTATGCTGCTGATCCA 840
QY 841 GACTGGATTCCTTTTATACATAGCAACGATATTTGGATATCTAAATGTAACACAGAA 900

Db 841 GACTGATTGCTTTTATACATAGCAACGATATTTTGGATATCTAAACATCGTAACCAGAGAA 900
Qy 901 GAAAGGAGACTCACATTATGTCGCACAATGAGCTAGCCAAACATGGAAGAATGCCAGATCA 960
Db 901 GAAAGGAGACTCACATTATGTCGCACAATGAGCTAGCCAAACATGGAAGAATGCCAGATCA 960
Qy 961 GCTGAGCTCGCTACCTTTGTTCTCCAAAGAAAGATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGAGCTCGCTACCTTTGTTCTCCAAAGAAAGATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGAAATTAATTCATGTTTACATCCCTATGTTGGAACAAAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGAAATTAATTCATGTTTACATCCCTATGTTGGAACAAAGGAGG 1140
Qy 1141 GCAGATTCAATTCCTGTTATCCCTTAAACAGGTACAGCAAACTCCTAAAGTCACTTTTAAGATG 1200
Db 1141 GCAGATTCAATTCCTGTTATCCCTTAAACAGGTACAGCAAACTCCTAAAGTCACTTTTAAGATG 1200
Qy 1201 TCAGAAATAATGATTTGATGCTGGAAGGAGGATCATAGATGTCATAGATAAGGAATTAATTT 1260
Db 1201 TCAGAAATAATGATTTGATGCTGGAAGGAGGATCATAGATGTCATAGATAAGGAATTAATTT 1260
Qy 1261 CAACCTTTTGAGATTTCTATTTTGAAGGAGTTGAAATATATTGCCAGAGCTGGATGACTCCT 1320
Db 1261 CAACCTTTTGAGATTTCTATTTGAGGAGTTGAAATATATTGCCAGAGCTGGATGACTCCT 1320
Qy 1321 GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Qy 1381 TTGATCTCACCTGAAATTAATTTATCCAGTGAAGATGATGTTATGAAAAGGAGAGACTC 1440
Db 1381 TTGATCTCACCTGAAATTAATTTATCCAGTGAAGATGATGTTATGAAAAGGAGAGACTC 1440
Qy 1441 ATTGAGTCAGTGCCTGATTTCTGTCAGCCCACTAATTAATCTATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATTTCTGTCAGCCCACTAATTAATCTATGAAGAAACACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTTATGTTTTTCCCAAGTCAAGAGAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTATGTTTTTCCCAAGTCAAGAGAGGAAATTTGAG 1560
Qy 1561 TTTATTTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCAITTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCAITTTATACAAAATTTACATCTATT 1620
Qy 1621 TTAAAGGAAAGCAATATAACGATCCAGTGGTGGCTGCTCCTCCAAAGTCAATTTCAAG 1680
Db 1621 TTAAAGGAAAGCAATATAACGATCCAGTGGTGGCTGCTCCTCCAAAGTCAATTTCAAG 1680
Qy 1681 TGCTCTATCAAGGAGATAGCAATTAACAGTGGTGGATGGGAGTTCTTGGCCGGCAT 1740
Db 1681 TGCTCTATCAAGGAGATAGCAATTAACAGTGGTGGATGGGAGTTCTTGGCCGGCAT 1740
Qy 1741 GGAATCAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGACCAAGAC 1800
Db 1741 GGAATCAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGACCAAGAC 1800
Qy 1801 TCCCTTTTGAAGCATACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGTGACAAGG 1860
Db 1801 TCCCTTTTGAAGCATACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACTTTGCTGTCAGTCAGCAGCTGTCACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACTTTGCTGTCAGTCAGCAGCTGTCACTTCTTTATA 1920
Qy 1921 AGTAAGTAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCT 1980
Db 1921 AGTAAGTAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCT 1980

Qy 1981 GAAGATGACCCAACTTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT 2040
Db 1981 GAAGATGACCCAACTTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTTC 2034
Qy 2041 CCTCTTCCTGACTACTACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGATTACA 2100
Db 2035 ----- 2034
Qy 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTG 2160
Db 2035 ----- 2034
Qy 2161 CTGTTCAATATATGGTGGTCTCAGGTGAGTGGTGAATTAATCTCGTTTAAAGAGTCAAG 2220
Db 2035 -----GTCCTCAGGTGAGTGGTGAATTAATCTCGTTTAAAGAGTCAAG 2078
Qy 2221 TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTGGTTGTAGTAGACAACAGG 2280
Db 2079 TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTGGTTGTAGTAGACAACAGG 2138
Qy 2281 GGATCTGTGTCACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA 2340
Db 2139 GGATCTGTGTCACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA 2198
Qy 2341 GAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2199 GAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2258
Qy 2401 TTAGATCTGTGGGATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTA 2460
Db 2259 TTAGATCTGTGGGATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTA 2318
Qy 2461 ATGCAGAGTCAAGATATCTTCAGGGTGGTATTCCTGGGCCCGCAGTCACTCTGTGGATC 2520
Db 2319 ATGCAGAGTCAAGATATCTTCAGGGTGGTATTCCTGGGCCCGCAGTCACTCTGTGGATC 2378
Qy 2521 TTCTATGATACAGGATACACGGAAGTTATATGGGTCACTCTGACAGAAATGAACAGGGC 2580
Db 2379 TTCTATGATACAGGATACACGGAAGTTATATGGGTCACTCTGACAGAAATGAACAGGGC 2438
Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATTCGTTTA 2640
Db 2439 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATTCGTTTA 2498
Qy 2641 CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGACATACAGTATATTAATCTG 2700
Db 2499 CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGACATACAGTATATTAATCTG 2558
Qy 2701 AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 2760
Db 2559 AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 2618
Qy 2761 AGCATAAGAGTTCTCTGAATCGGAGAACATTAATGAATCTTTTGGCACTACCTTCAA 2820
Db 2619 AGCATAAGAGTTCTCTGAATCGGAGAACATTAATGAATCTTTTGGCACTACCTTCAA 2678
Qy 2821 GAAAACCTTGGATCAGGTATGCTCTTAAAGTGATATAATTTTGCCTGTGTAGAAC 2880
Db 2679 GAAAACCTTGGATCAGGTATGCTCTTAAAGTGATATAATTTTGCCTGTGTAGAAC 2738
Qy 2881 TCTCTGGTATACATCTGGCTATTTTAAACAAATGAGAGGTTTAAATCAACAGAAAACACAGA 2940
Db 2739 TCTCTGGTATACATCTGGCTATTTTAAACAAATGAGAGGTTTAAATCAACAGAAAACACAGA 2798
Qy 2941 ATTGATCAATCAATTTTGATACCTGGCCATGTAAACATCTCTCTGAAAATAAATGTGGTG 3000
Db 2799 ATTGATCAATCAATTTTGATACCTGGCCATGTAAACATCTCTCTGAAAATAAATGTGGTG 2858
Qy 3001 CCATGACGGGTCTACGGTTTGTGGTAGTAATCTAAATCTTAATACCTTAAACCCCAATGCTCAAAA 3060
Db 2859 CCATGACGGGTCTACGGTTTGTGGTAGTAATCTAAATCTTAATACCTTAAACCCCAATGCTCAAAA 2918


```
QY 1201 TCAGAAATAATGATTGATGCTGAAGGAGGATCATAGATGCTATAGATAAGGAATTAATT 1260
Db 1201 TCAGAAATAATGATTGATGCTGAAGGAGGATCATAGATGCTATAGATAAGGAATTAATT 1260
QY 1261 CAACCTTTTGAGATCTTATTTGAAGGAGTTGAAATATATTCGCCAGACTGGATGGACTCCT 1320
Db 1261 CAACCTTTTGAGATCTTATTTGAAGGAGTTGAAATATATTCGCCAGACTGGATGGACTCCT 1320
QY 1321 GAGGAAATAATGCTTGCTCATCTACTAGATGCTCCAGACTCGCCCTACAGATAGTG 1380
Db 1321 GAGGAAATAATGCTTGCTCATCTACTAGATGCTCCAGACTCGCCCTACAGATAGTG 1380
QY 1381 TTGATCTCACTCGAATATTTATCCAGTGAAGATGATGTTATGAAAAGGAGAGACTC 1440
Db 1381 TTGATCTCACTCGAATATTTATCCAGTGAAGATGATGTTATGAAAAGGAGAGACTC 1440
QY 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCCACTAATATCTATGAAGAAACAAACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCCACTAATATCTATGAAGAAACAAACAGACATC 1500
QY 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCAACGAAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCAACGAAGGAAATTTGAG 1560
QY 1561 TTTATTTTGGCTCTGAAATCAAAACAGGTTTCGGTCATTTATACAAAATTAACATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAAATCAAAACAGGTTTCGGTCATTTATACAAAATTAACATCTATT 1620
QY 1621 TTAAGGAAGCAATATAACGATCCAGTGGTGGCTGCCTCCCAAGTGAATTTCAAG 1680
Db 1621 TTAAGGAAGCAATATAACGATCCAGTGGTGGCTGCCTCCCAAGTGAATTTCAAG 1680
QY 1681 TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db 1681 TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
QY 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
QY 1801 TCCCTTTTAGAGCATCACTGTACGTAGTCAGTTAAGTCACTGAGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACTGTACGTAGTCAGTTAAGTCACTGAGAGAGGTGACAAGG 1860
QY 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
QY 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTT 1980
QY 1981 GAGATGACCAACTTGCMAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
Db 1981 GAGATGACCAACTTGCMAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
QY 2041 CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGGATTTACA 2100
Db 2041 CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGGATTTACA 2100
QY 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGAAAGAAATATCCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
QY 2161 CTGTTTCATATATGTTGGTCTCTCAGGTGTCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTTCATATATGTTGGTCTCTCAGGTGTCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
QY 2221 TATTTCCGCTTTGAATACCCTAGCCCTCTCTAGGTTATGTGGTTATGTAGTAGACACAGG 2280
Db 2180 ----- 2179
Human; protein modification and maintenance molecule; PMM; cancer;
cell proliferation disorder; atherosclerosis; neurological disorder;
epilepsy; Huntington's disease; stroke; immune disorder; allergy;
inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
```

```
QY 2281 GGATCCTGTCACGAGGCGCTTAAATTTGAAGCGCGCTTTAAATATATAAAATGGGTCAAATA 2340
Db 2180 -----GGGTCAAATA 2189
QY 2341 GAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2190 GAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2349
QY 2401 TTAGATCGTGTGGGATCAACGGCTGGTCTTATGAGAGTACCTCTCCCTGATGGCATTA 2460
Db 2250 TTAGATCGTGTGGGATCAACGGCTGGTCTTATGAGAGTACCTCTCCCTGATGGCATTA 2309
QY 2461 ATGAGAGTCTAGATATCTTTCAGGGTGTCTATTCTGGGGCCCAAGTCACTCTGTGGATC 2520
Db 2310 ATGAGAGTCTAGATATCTTTCAGGGTGTCTATTCTGGGGCCCAAGTCACTCTGTGGATC 2369
QY 2521 TTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACCCAGAAATGAACAGGGC 2580
Db 2370 TTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACCCAGAAATGAACAGGGC 2429
QY 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA 2640
Db 2430 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA 2489
QY 2641 CTGCTCTTACATGGTTTCTCTGATGAGAAATGTCATTTTGACATACCACTATATTACTG 2700
Db 2490 CTGCTCTTACATGGTTTCTCTGATGAGAAATGTCATTTTGACATACCACTATATTACTG 2549
QY 2701 AGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAGACAC 2760
Db 2550 AGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAGACAC 2609
QY 2761 AGCATAAGAGTTCTTGAATCGGAGAACATTAAGCTGCATCTTTTGACTACCTTCAA 2820
Db 2610 AGCATAAGAGTTCTTGAATCGGAGAACATTAAGCTGCATCTTTTGACTACCTTCAA 2869
QY 2821 GAAAACCTTGGATACGTTATGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
Db 2670 GAAAACCTTGGATACGTTATGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAAC 2729
QY 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGA 2940
Db 2730 TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGA 2789
QY 2941 ATTGATCATCATTTTGATACCTGCCATGTAAACATCTACTCTGAAAAATAAATGTGGTG 3000
Db 2790 ATTGATCATCATTTTGATACCTGCCATGTAAACATCTACTCTGAAAAATAAATGTGGTG 2849
QY 3001 CCATGACGGGTCTACGGTTTGTGTAGTAAATCTAACTTAACCCCAATGCTCAAAA 3060
Db 2850 CCATGACGGGTCTACGGTTTGTGTAGTAAATCTAACTTAACCCCAATGCTCAAAA 2909
QY 3061 TCAAAATGATACATATTCCTGAGAGAGCCAGCAATACCATAAGATTAATAAAAAAAA 3120
Db 2910 TCAAAATGATACATATTCCTGAGAGAGCCAGCAATACCATAAGATTAATAAAAAAAA 2969
```

RESULT 8

ACA92421 standard; DNA; 2952 BP.

XX ACA92421;

AC ACA92421;

DT 15-JUL-2003 (first entry)

XX DNA encoding human PMM-6.

XX Human; protein modification and maintenance molecule; PMM; cancer;

KW cell proliferation disorder; atherosclerosis; neurological disorder;

KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;

KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;

KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;

KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;

KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thymimetic; gene; ds.
XX Homo sapiens.
XX WO2003031939-A2.
XX 17-APR-2003.
XX 11-OCT-2002; 2002WO-US032850.
XX 12-OCT-2001; 2001US-0329689P.
XX 25-OCT-2001; 2001US-0335703P.
XX 09-NOV-2001; 2001US-0348887P.
XX 28-NOV-2001; 2001US-0334145P.
XX 06-DEC-2001; 2001US-0337451P.
XX 14-DEC-2001; 2001US-0340584P.
XX (INCY-) INCYTE GENOMICS INC.
XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hatalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI; 2003-430274/40.
XX P-PSDB; ABU92026.
XX New human protein modification and maintenance molecules (PMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX Claim 5; Page 285-286; 31pp; English.
XX The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMM polypeptides (designated PMM
CC -1 to PMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
XX ACA92416-ACA92455 encode the human PMM polypeptides of the invention
XX
XX Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;
Query Match 89.3%; Score 2785.8; DB 10; Length 2952;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 2950; Conservative 0; Mismatches 2; Indels 153; Gaps 1;
QY 11 GCCTCGAGGCCAAGCGCGTCTACTGCGCGCGCTGCTTCTAGTCCGCTTCGCGC 70
DB 1 GCCTCGAGGCCAAGCGCGTCTACTGCGCGCGCTGCTTCTAGTCCGCTTCGCGC 60
QY 71 CTGGGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGACCGGAGTGGAGGC 130
DB 61 CTGGGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGACCGGAGTGGAGGC 120
QY 131 GCGCGAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGTCGCGCGCGG 190
DB 121 GCGCGAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGTCGCGCGCGG 180

QY 191 GCCGGGGGAAGGAAATGCAACATGCGCAGCAATGGAACAGAAACAGCTGGGTGTTG 250
DB 181 GCCGGGGGAAGGAAATGCAACATGCGCAGCAATGGAACAGAAACAGCTGGGTGTTG 240
QY 251 AGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCCTAAATTTG 310
DB 241 AGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCCTAAATTTG 300
QY 311 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAGAGCTGCTTGGCGATACCA 370
DB 301 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAGAGCTGCTTGGCGATACCA 360
QY 371 GAAATATACATGCTACATGATGCTAGGCAACACATGATTTTCATGTTTGTGAAGAGGA 430
DB 361 GAAATATACATGCTACATGATGCTAGGCAACACATGATTTTCATGTTTGTGAAGAGGA 420
QY 431 ATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTCGGTGAGAAC 490
DB 421 ATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTCGGTGAGAAC 480
QY 491 GAGAAATACACTGTTTATTTCTGAAATTCCTCAAAATCTATCAATAGAGAGCAGTCTTAA 550
DB 481 GAGAAATACACTGTTTATTTCTGAAATTCCTCAAAATCTATCAATAGAGAGCAGTCTTAA 540
QY 551 TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGACATATGGAATGTATT 610
DB 541 TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGACATATGGAATGTATT 600
QY 611 CTCGAGAGCAAGAACTATTAGAGAGAAAGAAACGATTTGGAAACAGTCCGAAATTCCTTCT 670
DB 601 CTCGAGAGCAAGAACTATTAGAGAGAAAGAAACGATTTGGAAACAGTCCGAAATTCCTTCT 660
QY 671 ACGATTATCACCAAGGAAAGTGGAACTTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGC 730
DB 661 ACGATTATCACCAAGGAAAGTGGAACTTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGC 720
QY 731 TAAAGATGGAGGGCCACAGGATTTAGCAACACCTTTTAAGCCCAATCTAGTGGAAA 790
DB 721 TAAAGATGGAGGGCCACAGGATTTAGCAACACCTTTTAAGCCCAATCTAGTGGAAA 780
QY 791 CTAGTTGTCCTCAACATACGATGGAATCAAAATTTATGCCCGCTGATCCAGACTGGATTG 850
DB 781 CTAGTTGTCCTCAACATACGATGGAATCAAAATTTATGCCCGCTGATCCAGACTGGATTG 840
QY 851 CTTTATATACATAGCAACGATTTTGGATATCTAACATCGTAAACAGAGAGAAAGAGAC 910
DB 841 CTTTATATACATAGCAACGATTTTGGATATCTAACATCGTAAACAGAGAGAAAGAGAC 900
QY 911 TCATTTATGTGCAATAGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCG 970
DB 901 TCATTTATGTGCAATAGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCG 960
QY 971 CTACCTTTGTTCTCCAGAGAAATTTGATAGATATCTTGGCTATTGTTGGTGGTCCAAAG 1030
DB 961 CTACCTTTGTTCTCCAGAGAAATTTGATAGATATCTTGGCTATTGTTGGTGGTCCAAAG 1020
QY 1031 CTGAAACAACTCCCGCTGGTGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGAAT 1090
DB 1021 CTGAAACAACTCCCGCTGGTGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGAAT 1080
QY 1091 CTGAGGTGGAATTTATTTCAATGTTACATCCCTTATGTTGGAACAAAGAGGGGAGATTCAT 1150
DB 1081 CTGAGGTGGAATTTATTTCAATGTTACATCCCTTATGTTGGAACAAAGAGGGGAGATTCAT 1140
QY 1151 TCGTGTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTGCAAAATPA 1210
DB 1141 TCGTGTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTGCAAAATPA 1200
QY 1211 TGATTGATGCTGAAGAGGATCATAGATGTCTATAGATAAGGAACTAATTTCAACCTTTTG 1270
DB 1201 TGATTGATGCTGAAGAGGATCATAGATGTCTATAGATAAGGAACTAATTTCAACCTTTTG 1260
QY 1271 AGATTCTTATTTGAAGGAGTTGAATATATTGCGCAGAGTGGATGGAGCTCTCTCTGAGGAAAT 1330

Db 1261 AGATTCTATTGAGGAGTTGAAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAAAT 1320
Qy 1331 ATGCTTGGTCCATCCTACTAGATCGCTCCAGATCGCTCAGATAGTGTGATCTCAC 1390
Db 1321 ATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCCGCTCAGATAGTGTGATCTCAC 1380
Qy 1391 CTGAATATTTATCCGAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAG 1450
Db 1381 CTGAATATTTATCCGAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAG 1440
Qy 1451 TGCCTGATCTGTGACGCCACTAATTATCTATGAAGAAACAAAGACATCTGGAATAATA 1510
Db 1441 TGCCTGATCTGTGACGCCACTAATTATCTATGAAGAAACAAAGACATCTGGAATAATA 1500
Qy 1511 TCCATGACATCTTTTCATGTTTTTCCCAAGTCCAGAGAGGAAATTTGAGTTTATTTTTG 1570
Db 1501 TCCATGACATCTTTTCATGTTTTTCCCAAGTCCAGAGAGGAAATTTGAGTTTATTTTTG 1560
Qy 1571 CCTCTGAATGCAAAACAGGTTTCCGTCAITTTATACAAAATTACATCTATTTTAAAGGAAA 1630
Db 1561 CCTCTGAATGCAAAACAGGTTTCCGTCAITTTATACAAAATTACATCTATTTTAAAGGAAA 1620
Qy 1631 GCAAAATATAACGATCCAGTGGTGGCTGCTCCCAAGTGATTTCAAGTGTCTATCA 1690
Db 1621 GCAAAATATAACGATCCAGTGGTGGCTGCTCCCAAGTGATTTCAAGTGTCTATCA 1680
Qy 1691 AAGAGGATAGCAATTTACCAGTGGTGAATGGGAAGTCTTGGCGGCATGGATCTAATA 1750
Db 1681 AAGAGGATAGCAATTTACCAGTGGTGAATGGGAAGTCTTGGCGGCATGGATCTAATA 1740
Qy 1751 TCCAAGTTGATGAAGTCCAGAGGCTGGTATTTTGAAGGCCCAAGAGCTCCCTTTTAC 1810
Db 1741 TCCAAGTTGATGAAGTCCAGAGGCTGGTATTTTGAAGGCCCAAGAGCTCCCTTTTAC 1800
Qy 1811 AGCATACCTGTAGTGTAGTCAAGTACGTAACTCTGGAGAGGTGAAGGCTGATGACC 1870
Db 1801 AGCATACCTGTAGTGTAGTCAAGTACGTAACTCTGGAGAGGTGAAGGCTGATGACC 1860
Qy 1931 GTGGCTACTCAATCTTGTGTGATCAGTCAAGTGTGACCTCTTTTATAAGTAAAGTATA 1930
Db 1861 GTGGCTACTCAATCTTGTGTGATCAGTCAAGTGTGACCTCTTTTATAAGTAAAGTATA 1920
Qy 1931 GTAAACCAAGAAATCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACC 1990
Db 1921 GTAAACCAAGAAATCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACC 1980
Qy 1991 CAACTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCCTG 2050
Db 1981 CAACTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCCTG 2040
Qy 2051 ACTATACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGA 2110
Db 2041 ACTATACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGA 2100
Qy 2111 TGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGTGTTTCATAT 2170
Db 2101 TGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGTGTTTCATAT 2160
Qy 2171 ATGCTGTGCTCAGTGTGAGTGTGTAATTCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230
Db 2161 ATGCTGTGCTCAGTGTGAGTGTGTAATTCGGTTTAAAGGAGTCAAGTATTTCCGCT 2220
Qy 2231 TGAATACCTTAGCCTCTCTAGGTTATGTGGTTAGTGTATAGACAAACAGGGGATCCTGTC 2290
Db 2221 TGAATACCTTAGCCTCTCTAGGTTATGTGGTTAGTGTATAGACAAACAGGGGATCCTGTC 2280
Qy 2291 ACCGAGGGCTTTAAATTTGAAAGGCCCTTTTAAATATATAAATATGAAATATGACG 2350
Db 2281 ACCGAGGGCTTTAAATTTGAAAGGCCCTTTTAAATATATAAAT----- 2320
Qy 2351 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTTAGATCGTG 2410

Db 2321 ----- 2320
Qy 2411 TGGGCATCCAGGCTGGTCTCTATGAGGAGTACCTCTCCCTGATGGCATTAATGACAGGT 2470
Db 2321 ----- 2320
Qy 2471 CAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATA 2530
Db 2321 -----GTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATA 2367
Qy 2531 CAGGATACACGGAAGCTTTATATGGGTCACTCTGACCGAATGAACAGGGCTATTACTTAG 2590
Db 2368 CAGGATACACGGAAGCTTTATATGGGTCACTCTGACCGAATGAACAGGGCTATTACTTAG 2427
Qy 2591 GATCTGTGGCATGCAAGCAGAAAGTTCCCTCTCGAACCAAAATCGTTTACTGTCTTTAC 2650
Db 2428 GATCTGTGGCATGCAAGCAGAAAGTTCCCTCTCGAACCAAAATCGTTTACTGTCTTTAC 2487
Qy 2651 ATGGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCACTATATTAATCTGAGTTTTTATAG 2710
Db 2488 ATGGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCACTATATTAATCTGAGTTTTTATAG 2547
Qy 2711 TGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAG 2770
Db 2548 TGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAG 2607
Qy 2771 TTCTCTGAATCCGGAGAACATTATGAACCTGTCATCTTTTGCACACTACCTTCAAGAAAACTTG 2830
Db 2608 TTCTCTGAATCTGGAGAACATTATGAACCTGTCATCTTTTGCACACTACCTTCAAGAAAACTTG 2667
Qy 2831 GATCAGTATTGCTGCTCTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTCGTAT 2890
Db 2668 GATCAGTATTGCTGCTCTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTCGTAT 2727
Qy 2891 ACACGTGGCTATTTAACCAATGAGAGGTTTAACTACACAGAAACACAGAAATGATCATC 2950
Db 2728 ACACGTGGCTATTTAACCAATGAGAGGTTTAACTACACAGAAACACAGAAATGATCATC 2787
Qy 2951 ACATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAAATGTGGTCCCATGAGGG 3010
Db 2788 ACATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAAATGTGGTCCCATGAGGG 2847
Qy 3011 GTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCCATGCTCTCAAAATCAAAATGATA 3070
Db 2848 GTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCCATGCTCTCAAAATCAAAATGATA 2907
Qy 3071 CATATTCTTGAGAGACCCAGCAATACCATTAAGAAATTAATAAAAAA 3115
Db 2908 CATATTCTTGAGAGACCCAGCAATACCATTAAGAAATTAATAAAAAA 2952

RESULT 9
ADL13374
ID ADL13374 standard; cDNA; 2797 BP.
XX
AC ADL13374;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human steroid-induced C3A liver cell cDNA #1103.
XX
KW ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
XX steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
OS Homo sapiens.
XX
PN US6673549-B1.
XX
PD 06-JAN-2004.
XX
PF 12-OCT-2001; 2001US-00976594.
XX
PR 12-OCT-2000; 2000US-0240409P.

XX (INCY-) INCYTE CORP.
XX PA
XX PI Furness LM, Buchbinder JL;
XX DR WPI; 2004-068610/07.
XX PT Combination useful for preparing a composition for treating liver
XX disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
XX PT comprises cDNAs that are differentially expressed in response to steroid
XX treatment.
XX PS Claim 1; SEQ ID NO 1103; 141pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;
Query Match 88.4%; Score 2759.4; DB 12; Length 2797;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;
QY 335 CCTGGAGTCAGCTTAAAGAGCTGCTGGCGATACAGAGAAATATCATGGCTACATGATGG 394
DB 1 CCTGGAGTCAGCTTAAAGAGCTGCTGGCGATACAGAGAAATATCATGGCTACATGATGG 60
QY 395 CTAAGGACACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 454
DB 61 CTAAGGACACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 120
QY 455 ACAGAACTATTACCTTGGCCATGCTGGTGAGAGACAGAGAAATACATGTTTATTCTG 514
DB 121 ACAGAACTATTACCTTGGCCATGCTGGTGAGAGACAGAGAAATACATGTTTATTCTG 180
QY 515 AAATTCGCAAACTATCAATAGAGACAGCTTAAATGCTCTCTTGGAGGCTCTTTTGG 574
DB 181 AAATTCGCAAACTATCAATAGAGACAGCTTAAATGCTCTCTTGGAGGCTCTTTTGG 240
QY 575 ATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGAGAACTATTAAAG 634
DB 241 ATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGAGAACTATTAAAG 300
QY 635 AAAGAAACGATTTGGAACAGCTGGAAATGCTTTTACGATTTATCACAAGAGAGTGGAA 694
DB 301 AAAGAAACGATTTGGAACAGCTGGAAATGCTTTTACGATTTATCACCAGAGAGTGGAA 360
QY 695 CATTTCTCTTTCAAGCCGCTAGTGGAAATTTATCACGCTAAAGAGATGGAGGCGCCACAAGGAT 754
DB 361 CATTTCTCTTTCAAGCCGCTAGTGGAAATTTATCACGCTAAAGAGATGGAGGCGCCACAAGGAT 420
QY 755 TTACGCAACAACTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATGG 814
DB 421 TTACGCAACAACTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATGG 480
QY 815 ATCCAAATATATGCGCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTT 874

DB 481 ATCCAAATATATGCGCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTT 540
QY 875 GGATATCTAATCGTAAACAGAGAGAGAGAGAGACTCATTATGTGACAAATGAGCTAG 934
DB 541 GGATATCTAATCGTAAACAGAGAGAGAGAGAGACTCATTATGTGACAAATGAGCTAG 600
QY 935 CCAACATGGAAGAGATGCGCAGATCAGCTGGAGTGGCTACCTTTTGTCTTCCAAAGAGAT 994
DB 601 CCAACATGGAAGAGATGCGCAGATCAGCTGGAGTGGCTACCTTTTGTCTTCCAAAGAGAT 660
QY 995 TTGATAGATATCTGGCTATTTGGTGTCTCCAAAGAGTGAACAACTCCAGTGGTGTA 1054
DB 661 TTGATAGATATCTGGCTATTTGGTGTCTCCAAAGAGTGAACAACTCCAGTGGTGTA 720
QY 1055 AAATTTTAGAATCTTATATGAAGAAATGATGATCTGAGTGGAAATTTATCATGTTA 1114
DB 721 AAATTTTAGAATCTTATATGAAGAAATGATGATCTGAGTGGAAATTTATCATGTTA 780
QY 1115 CATCCCCATGTTGGAACCAAGG-AGGCGAGATTCATTCCGTTATCCTTAAACAGGATACA 1173
DB 781 CATCCCCATGTTGGAACCAAGGCGAGGAGGAGATTCATTCCGTTATCCTTAAACAGGATACA 840
QY 1174 GCAATCTTAAGTCACTTTTAAAGATGTCAGAAATAATGATTTGATGCTGAAAGAGGATC 1233
DB 841 GCAATCTTAAGTCACTTTTAAAGATGTCAGAAATAATGATTTGATGCTGAAAGAGGATC 900
QY 1234 ATAGATGCTATAGATAAGGAATTAATCAACCTTTTGAAGATTCATTGAAAGAGGATGAA 1293
DB 901 ATAGATGCTATAGATAAGGAATTAATCAACCTTTTGAAGATTCATTGAAAGAGGATGAA 960
QY 1294 TATATTCGCGAGAGCTGGATGGACTCCTGAGGGAATA----TGCTTGGTCCATCTACT 1349
DB 961 TATATTCGCGAGAGCTGGATGGACTCCTGAGGGAATAATGCTGCTGCTGCTCATCTACT 1020
QY 1350 AGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAGTGAATTTATTTATCCAGT 1409
DB 1021 AGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAGTGAATTTATTTATCCAGT 1080
QY 1410 AGAAGATGATGTTATGGAAGGAGAGACTCATTGATGAGTGGCTGATCTGTGACGCC 1469
DB 1081 AGAAGATGATGTTATGGAAGGAGAGACTCATTGATGAGTGGCTGATCTGTGACGCC 1140
QY 1470 ACTAATTTATCTTGAAGAAACACACATCTGGATAAATATCCATGATCATCTTTCATGT 1529
DB 1141 ACTAATTTATCTTGAAGAAACACACATCTGGATAAATATCCATGATCATCTTTCATGT 1200
QY 1530 TTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGG 1589
DB 1201 TTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGG 1260
QY 1590 TTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAACAAATATAACGATCCAG 1649
DB 1261 TTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAACAAATATAACGATCCAG 1320
QY 1650 TGGTGGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGCAATAC 1709
DB 1321 TGGTGGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGCAATAC 1380
QY 1710 CAGTGGTGAATGGGAAGTCTTGGCGGCGATGATCTAATATATCAAGTTGATGAGTACG 1769
DB 1381 CAGTGGTGAATGGGAAGTCTTGGCGGCGATGATCTAATATATCAAGTTGATGAGTACG 1440
QY 1770 AAGGCTGTATATTTTGAAGGCAACAGGCTGCTCCCTTTAGAGCATCCTCTGATGATG 1829
DB 1441 AAGGCTGTATATTTTGAAGGCAACAGGCTGCTCCCTTTAGAGCATCCTCTGATGATG 1500
QY 1830 CAGTTACGTAATCTCTGAGAGGTTGCAAGGCTGATGACCGTGGCTACTCACAATCTTG 1889
DB 1501 CAGTTACGTAATCTCTGAGAGGTTGCAAGGCTGATGACCGTGGCTACTCACAATCTTG 1560
QY 1890 CTGCAATCAGTCCAGCTGCTGATCTTTTATAGTATAGTAACTAGTAACCAAGAGATCCACA 1949

Db 1561 CTGCATCAGTCAGCACTGTGACITCTTTTATAGTAAGTATAGTAAACCAAGAAATCCACA 1620
Qy 1950 CTGTGTGTCCTTTTACAAAGCTATCAAGTCCTGAAGATGACCCAACTTGCAAAACAAGGA 2009
Db 1621 CTGTGTGTCCTTTTACAAAGCTATCAAGTCCTGAAGATGACCCAACTTGCAAAACAAGGA 1680
Qy 2010 ATTTTGGGCGACCACTTTTGGGATTCAGCAGGTCTCTTCTGCTGACTATACCTCTCCAGAAAT 2069
Db 1681 ATTTTGGGCGACCACTTTTGGGATTCAGCAGGTCTCTTCTGCTGACTATACCTCTCCAGAAAT 1740
Qy 2070 TTTCTCTTTTGAAGTACTACTGATTTTCAATTTGATGGATGCTCTACAAGCTTCATGA 2129
Db 1741 TTTCTCTTTTGAAGTACTACTGATTTTCAATTTGATGGATGCTCTACAAGCTTCATGA 1800
Qy 2130 TCTACAGCTGGAAAGAAATCTACTGTGCTGCTCATATATATGTTGGTCTCTCAGGTGCA 2189
Db 1801 TCTACAGCTGGAAAGAAATCTACTGTGCTGCTCATATATATGTTGGTCTCTCAGGTGCA 1860
Qy 2190 GTTGTGTAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGATACCTAGCCTCTCT 2249
Db 1861 GTTGTGTAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGATACCTAGCCTCTCT 1920
Qy 2250 AGGTTATGTGTTGATGATAGACAAACAGGGGATCTGTACCGAGGGCTTAAATTTGA 2309
Db 1921 AGGTTATGTGTTGATGATAGACAAACAGGGGATCTGTACCGAGGGCTTAAATTTGA 1980
Qy 2310 AGGCGCTTTAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAAGGACTCA 2369
Db 1981 AGGCGCTTTAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAAGGACTCA 2040
Qy 2370 ATATCTAGCTTCTCGATATGATTTCAITTTAGCTTAGATCGTGTGGGATCCACGGCTGGTC 2429
Db 2041 ATATCTAGCTTCTCGATATGATTTCAITTTAGCTTAGATCGTGTGGGATCCACGGCTGGTC 2100
Qy 2430 CTATGAGGATACCTCTCCCTGATGGCAATTAATCAGAGGTGAGATATCTTCAGGTTGC 2489
Db 2101 CTATGAGGATACCTCTCCCTGATGGCAATTAATCAGAGGTGAGATATCTTCAGGTTGC 2160
Qy 2490 TATTGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGTTA 2549
Db 2161 TATTGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGTTA 2220
Qy 2550 TATGGGTCACTCTGACCGAGATGACAGGGCTATTACTAGGATCTGTGGCCATGCAAGC 2609
Db 2221 TATGGGTCACTCTGACCGAGATGACAGGGCTATTACTAGGATCTGTGGCCATGCAAGC 2280
Qy 2610 AGAAAAGTTCCCTCTGAAACCAATCGTTTACTCTCTTACATGGTTTCTCGATGAGAA 2669
Db 2281 AGAAAAGTTCCCTCTGAAACCAATCGTTTACTCTCTTACATGGTTTCTCGATGAGAA 2340
Qy 2670 TGTCCTATTTGCACATACAGTATATTAATCTGAGTTTATAGTAGGGCTGGAAAGCCATA 2729
Db 2341 TGTCCTATTTGCACATACAGTATATTAATCTGAGTTTATAGTAGGGCTGGAAAGCCATA 2400
Qy 2730 TGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCGAATTCGGGAGAA 2789
Db 2401 TGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCGAATTCGGGAGAA 2460
Qy 2790 TTATGAATGCACTCTTTTGCATCTACCTTCAAGAAACCTTTGGATCAGTATTCCTGCTCT 2849
Db 2461 TTATGAATGCACTCTTTTGCATCTACCTTCAAGAAACCTTTGGATCAGTATTCCTGCTCT 2520
Qy 2850 AAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAA 2909
Db 2521 AAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAA 2580
Qy 2910 ATGAGGAGTTTAAATCAACAGAAACACAGAAATTCATCATCAATTTTGTATCTGCGAT 2969
Db 2581 ATGAGGAGTTTAAATCAACAGAAACACAGAAATTCATCATCAATTTTGTATCTGCGAT 2640
Qy 2970 GTACATCTACTCTCAAAATAAATGTGTGCGCATGACAGGGGTCTACGGTTTGTGGTAGT 3029
Db 2641 GTACATCTACTCTCAAAATAAATGTGTGCGCATGACAGGGGTCTACGGTTTGTGGTAGT 2700

Qy 3030 AATCTATATACCTTTAACCCCAATGCTCAAAATCAAAATGATACATATCTCTGAGACCCA 3089
Db 2701 AATCTATATACCTTTAACCCCAATGCTCAAAATCAAAATGATACATATCTCTGAGACCCA 2760
Qy 3090 GCAATACCAATGAATTTACTTAAAAA 3120
Db 2761 GCAATACCAATGAATTTACTAAAAA 2791
RESULT 10
ACA92425
ID ACA92425 standard; DNA; 2929 BP.
XX
AC ACA92425;
DT 15-JUL-2003 (first entry)
XX
DE DNA encoding human PMMM-10.
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
cell proliferation disorder; atherosclerosis; neurological disorder;
epilepsy; Huntington's disease; stroke; immune disorder; allergy;
inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
antiinflammatory; thymolaminetic; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
WI: 2003-430274/40.
DR P-PSDB; AB092030.
XX
PT New human protein modification and maintenance molecules (PMMM), useful
for diagnosing, treating and preventing diseases or conditions associated
with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
infections.
XX
PS Claim 5; Page 289; 31pp; English.
XX
CC The present invention relates to the isolation of human protein
modification and maintenance molecules (PMMM), and the polynucleotide
sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
-1 to PMMM-40) are disclosed. The sequences of the invention are useful
for diagnosing a condition or disease associated with the expression of
PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
generating an expression profile of a sample containing the
polynucleotides. The diseases or conditions associated with decreased
expression or overexpression of PMMM are cell proliferation disorders
(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,

CC	allergies), developmental disorders (e.g. hypothyroidism, Cushing's syndrome), gastrointestinal or epithelial disorders, and infections. The PMM polymorphisms or their fragments are useful in screening compounds for effectiveness as agonists or antagonists of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to, or modulate the activity of the polypeptide.
CC	ACA92416-ACA92455 encode the human PMM polypeptides of the invention
XX	
SQ	Sequence 2929 BP; 864 A; 606 C; 661 G; 798 T; 0 U; 0 Other;
Query Match	87.9%; Score 2742.4; DB 10; Length 2929;
Best Local Similarity	94.4%; Pred. No. 0;
Matches 2927; Conservative	0; Mismatches 1; Indels 174; Gaps 1
QY	13 CTCGAGGCCAAGCGCGTCTACTATGCGCGCGTCTTCTTAGTGTGCGGTTCGCGCGCT 72
DB	2 CTCGAGGCCAAGCGCGTCTACTATGCGCGCGTCTTCTTAGTGTGCGGTTCGCGCGCT 61
QY	73 GGGTGTCTACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGCCGAGGTGGAGCGG 132
DB	62 GGGTGTCTACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGCCGAGGTGGAGCGG 121
QY	133 CGCAGCATGAAGCGCGCAGCGCGCTCCATAGCGCAGCTCGCGAGCGTCCGCGCGGGCG 192
DB	122 CGCAGCATGAAGCGCGCAGCGCGCTCCATAGCGCAGCTCGCGAGCGTCCGCGCGGGC 181
QY	193 CGGGGGGAAGAAAAATGCAACATGCGCAGCAATGGAAACAGAACAGCTGGGTGTTGAG 252
DB	182 CGGGGGGAAGAAAAATGCAACATGCGCAGCAATGGAAACAGAACAGCTGGGTGTTGAG 241
QY	253 ATATTGAAACTGCGGACTGTGAGAGAAATTGAATCAGAGATCGGCTTAATTGGAG 312
DB	242 ATATTGAAACTGCGGACTGTGAGAGAAATTGAATCAGAGATCGGCTTAATTGGAG 301
QY	313 CCTTTTATGTTGAGCGGTATTCTCTGGAGTCAGCTTAAAAAGCTGCTTGCAGATACCAG 372
DB	302 CCTTTTATGTTGAGCGGTATTCTCTGGAGTCAGCTTAAAAAGCTGCTTGCAGATACCAG 361
QY	373 AAATATCATGGCTACATGATGCTTAAGGCACACATGATTTCTGTTGTAAGAGGAAT 432
DB	362 AAATATCATGGCTACATGATGCTTAAGGCACACATGATTTCTGTTGTAAGAGGAAT 421
QY	433 GATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGGCATGTCTGTTGAGAACAG 492
DB	422 GATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGGCATGTCTGTTGAGAACAG 481
QY	493 GAAATACACTGTTTATTCTGAAATTCCTCAAACTATCAATAGACAGCAGCTTAAATG 552
DB	482 GAAATACACTGTTTATTCTGAAATTCCTCAAACTATCAATAGACAGCAGCTTAAATG 541
QY	553 CTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGAATATGGAATGTATCT 612
DB	542 CTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGAATATGGAATGTATCT 573
QY	613 CGAAGAAAGAACTATTAGAGAAAGAAAAACGCAATGGAACTGGAAATTTCTTCTTAC 672
DB	574 ----- 573
QY	673 GATTATCAACAAGGAAGTGGAACTTTCTGTTTCAAGCGGTAGTGGAAATTTATCATCGTA 732
DB	574 ----- 573
QY	733 AAAGATGGAGGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACT 792
DB	574 -----GCACAACTTTAAGGCCCAATCTAGTGGAACT 607
QY	793 AGTTGTCCCAACATACGATAGGATCCAAAATTTATGCCCCGTGATTCAGACTGGAATGCT 852
DB	608 AGTTGTCCCAACATACGATAGGATCCAAAATTTATGCCCTGCTGATCCAGACTGGAATGCT 667
QY	853 TTTATACATGCAACGATATTGGATATCTAACATCTGTAACACAGAGAAAGAGGACTC 912
DB	668 TTTATACATGCAACGATATTGGATATCTAACATCTGTAACCAAGAGAAAGAGGACTC 727

QY 1993 ACTTGCRAAACAAGGATTTTGGGCCACCATTTTGGATTCAGAGTGCTCTCTCTGCAC 2052
DB 1808 ACTTGCRAAACAAGGATTTTGGGCCACCATTTTGGATTCAGAGTGCTCTCTCTGCAC 1867
QY 2053 TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATGTATGGGATG 2112
DB 1868 TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATGTATGGGATG 1927
QY 2113 CTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGTCTGTTCATATAT 2172
DB 1928 CTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGTGTTCATATAT 1987
QY 2173 GGTGTCTCTCAGGTGCGATGATTAATCGGTTTAAAGGATCAAGTATTTCCGCTTG 2232
DB 1988 GGTGTCTCTCAGGTGCGATGATTAATCGGTTTAAAGGATCAAGTATTTCCGCTTG 2047
QY 2233 AATACCCTAGGCTCTCTAGGTTATGTGGTTGTAGTAGACACAACAGGGGATCCCTGTAC 2292
DB 2048 AATACCCTAGGCTCTCTAGGTTATGTGGTTGTAGTAGACACAACAGGGGATCCCTGTAC 2107
QY 2293 CGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAATGGGTCAAAATAGAAATTCAGCAT 2352
DB 2108 CGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAATGGGTCAAAATAGAAATTCAGCAT 2167
QY 2353 CAGGTGGAAGGACTCCAAATCTAGCTTCTCGATATGATTTCAATTGACCTTAGATCGTGTG 2412
DB 2168 CAGGTGGAAGGACTCCAAATCTAGCTTCTCGATATGATTTCAATTGACCTTAGATCGTGTG 2227
QY 2413 GGCATCCAGGCTGTCTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGGTCA 2472
DB 2228 GGCATCCAGGCTGTCTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGGTCA 2287
QY 2473 GATATCTTCAGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA 2532
DB 2288 GATATCTTCAGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA 2347
QY 2533 GGATACACGGAAGTTATATGGGTACCCCTGACAGAAATGAACAGGGCTTATTACTTAGGA 2592
DB 2348 GGATACACGGAAGTTATATGGGTACCCCTGACAGAAATGAACAGGGCTTATTACTTAGGA 2407
QY 2593 TCTGTGGCCATGCAAGCAGAAAGTTCCCTCTTGAAACCAATCGTTTACTGCTCTTACAT 2652
DB 2408 TCTGTGGCCATGCAAGCAGAAAGTTCCCTCTTGAAACCAATCGTTTACTGCTCTTACAT 2467
QY 2653 GGTTCCTCTGGATGGAATGTCCATTTTGGACATACAGATATATTACTGAGTTTTTTAGTG 2712
DB 2468 GGTTCCTCTGGATGGAATGTCCATTTTGGACATACAGATATATTACTGAGTTTTTTAGTG 2527
QY 2713 AGGGCTGAAAAGCCATATGATTTACAGATCTATCTCCTCAGGAGACACAGCATAAGAGTT 2772
DB 2528 AGGGCTGAAAAGCCATATGATTTACAGATCTATCTCCTCAGGAGACACAGCATAAGAGTT 2587
QY 2773 CCTGAATCGGAGACATTTATGACTGCATCTTTTGCACCTTACCTTCAAGAAACCTTGA 2832
DB 2588 CCTGAATCGGAGAACATTTATGAATGCACTCTTTTGCACCTTACCTTCAAGAAACCTTGA 2647
QY 2833 TCAGTATTTGCTCTTAAAGTGATATAATTTTGAACCTGTGTAGAACCTCTCTGGTATAC 2892
DB 2648 TCAGTATTTGCTCTTAAAGTGATATAATTTTGAACCTGTGTAGAACCTCTCTGGTATAC 2707
QY 2893 ACTGCTATTTAAACAAATGAGGAGTTTAAATCAACGAAACACAGAAATGATCATCAC 2952
DB 2708 ACTGCTATTTAAACAAATGAGGAGTTTAAATCAACGAAACACAGAAATGATCATCAC 2767
QY 2953 ATTTTGATACCTGCCATGTAAACATCTACTCTGAAATATAATGTGGTGCATGCGAGGGT 3012
DB 2768 ATTTTGATACCTGCCATGTAAACATCTACTCTGAAATATAATGTGGTGCATGCGAGGGT 2827
QY 3013 CTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCCAACATGCTCAAAATCAAATGATACA 3072
DB 2828 CTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCCAACATGCTCAAAATCAAATGATACA 2887
QY 3073 TATTCCTGAGAGACCCAGCAATACCAATAAGAATTTACTAAAAA 3114

DB 2888 TATTCCTGAGAGCCAGCAATACCAATAGAAATTTACTAAAAA 2929
RESULT 11
ADV43981
ID ADV43981 standard; cDNA; 2649 BP.
XX ADV43981;
DT 10-MAR-2005 (first entry)
XX Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 1609.
XX microarray; psychoneuroendocrinimmune; chronic fatigue;
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW cancer; neoplasm; infection; expressed sequence tag; ss.
XX Homo sapiens.
XX OS
XX WO2004108899-A2.
XX PN
XX 16-DEC-2004.
XX PD
XX 04-JUN-2004; 2004WO-US017686.
XX PF
XX 04-JUN-2003; 2003US-0475915P.
XX PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PA
XX Nicholson A, Vernon SD;
XX PI
XX WPI; 2005-031682/03.
XX DR
XX New microarray comprising probes for genes involved in
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT condition associated with PNI activity, e.g., inflammatory or infectious
PT diseases.
XX PS
XX Claim 1; SEQ ID NO 1609; 254pp; English.
XX CC
XX The invention relates to a new microarray which comprises probes for
CC genes involved in psychoneuroendocrinimmune (PNI) activity. The
CC microarray is useful in diagnosing a condition associated with PNI
CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
CC cancer and infection. The present sequence represents a
CC psychoneuroendocrinimmune gene expressed sequence tag. Note the
CC specificatio mentions SEQ ID NO of up to 3314 but only sequences up to
CC SEQ ID NO 1829 are provided.
XX SQ
XX Sequence 2649 BP; 804 A; 514 C; 585 G; 746 T; 0 U; 0 Other;
Query Match 84.9%; Score 2647.4; DB 14; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGTCTTGAGATATTTGAACTCGGACTGT 273
DB 1 ATGGCAGCAGCAATCGAAACAGACAGCTGGTCTTGAGATATTTGAACTCGGACTGT 60
QY 274 GAGGAGATATTTGAATCAAGATCGGCTTAAATTTGGAGCTTTTATGTTGACCGGTAT 333
DB 61 GAGGAGATATTTGAATCAAGATCGGCTTAAATTTGGAGCTTTTATGTTGACCGGTAT 120
QY 334 TCTGTGAGTCACTTAAAGCTGCTTCCGATACAGAAATATCATGCTACATGATG 393
DB 121 TCTGTGAGTCACTTAAAGCTGCTTCCGATACAGAAATATCATGCTACATGATG 180
QY 394 GCTAAGGCCACCATGATTTTATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
DB 181 GCTAAGGCCACCATGATTTTATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 240
QY 454 GACAGAATCTATTACCTTTGCCATGTCTGGTGAGAACACAGAAAAATACACTGTTTTATTCT 513

Db	241	GACAGAACTATTACCTTTGCCATGCTGGTGAGAACAGAGAAATACATGTTTTATTCT	300
Qy	514	GAATTTCCAAATCATCAATAGACGACAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTG	573
Db	301	GAATTTCCAAATCATCAATAGACGACAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTG	360
Qy	574	GATCTTTTTCAGGCAACACCTGGACTATGGAATGATTTCTCGAGAAGAACTATTAAGA	633
Db	361	GATCTTTTTCAGGCAACACCTGGACTATGGAATGATTTCTCGAGAAGAACTATTAAGA	420
Qy	634	GAAGAAACCCATTTGGAACAGTCGGAATGCTTTCTTAACGATTATCACCAAGGAAGTGA	693
Db	421	GAAGAAACCCATTTGGAACAGTCGGAATGCTTTCTTAACGATTATCACCAAGGAAGTGA	480
Qy	694	ACATTTCTGTTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGGCCACAAGA	753
Db	481	ACATTTCTGTTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGGCCACAAGA	540
Qy	754	TTTACGCAACAACTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGGATG	813
Db	541	TTTACGCAACAACTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGGATG	600
Qy	814	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATATACATAGCAACGATAT	873
Db	601	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATATACATAGCAACGATAT	660
Qy	874	TGGATATCTTAACATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCAACATAGCTTA	933
Db	661	TGGATATCTTAACATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCAACATAGCTTA	720
Qy	934	GCCAACTGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA	993
Db	721	GCCAACTGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA	780
Qy	994	TTTGTATAGATATCTGGCTATTTGGTGTGTCAAAAGCTGAAAACAACTCCGAGTGGTGT	1053
Db	781	TTTGTATAGATATCTGGCTATTTGGTGTGTCAAAAGCTGAAAACAACTCCGAGTGGTGT	840
Qy	1054	AAAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTAATCATGTT	1113
Db	841	AAAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTAATCATGTT	900
Qy	1114	ACATCCCTATGTTGGAACCAAGAGGGCGAGATTCATTCGGTTATCTCTAAACAGGTACA	1173
Db	901	ACATCCCTATGTTGGAACCAAGAGGGCGAGATTCATTCGGTTATCTCTAAACAGGTACA	960
Qy	1174	GCAATCTCTAAAGTCATTTTAAGATGTCAGAAATAATGATGCTGAAGAGGATC	1233
Db	961	GCAATCTCTAAAGTCATTTTAAGATGTCAGAAATAATGATGCTGAAGAGGATC	1020
Qy	1234	ATAGATGTCATAGATAAGAACTAATTCACCTTTTCAGATTCCTATTTGAAGAGTTGAA	1293
Db	1021	ATAGATGTCATAGATAAGAACTAATTCACCTTTTCAGATTCCTATTTGAAGAGTTGAA	1080
Qy	1294	TATATGTCAGAGCTGGATGACCTCTGAGGAAATATGCTTGGTCCATCTCTACTAGAT	1353
Db	1081	TATATGTCAGAGCTGGATGACCTCTGAGGAAATATGCTTGGTCCATCTCTACTAGAT	1140
Qy	1354	CGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAATTAATTTATCCAGTAGAA	1413
Db	1141	CGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAATTAATTTATCCAGTAGAA	1200
Qy	1414	GATGATGTTATGGAAGCGCAGAGACTCATTTGAGTCAGTCGCTGATTCGTGAGCCACTA	1473
Db	1201	GATGATGTTATGGAAGCGCAGAGACTCATTTGAGTCAGTCGCTGATTCGTGAGCCACTA	1260
Qy	1474	ATTATCTATGAGAAACACAGACATCTGGAATAATATCCATGACATCTTTCAATGTTTTT	1533
Db	1261	ATTATCTATGAGAAACACAGACATCTGGAATAATATCCATGACATCTTTCAATGTTTTT	1320
Qy	1534	CCCCAAGTTCAGAGAGGAAATTCAGTTTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTC	1593

Db	1321	CCCCAAGTTCAGAGAGGAAATTCAGTTTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTC	1380
Qy	1594	CGTCATTTATACAAAATTAATCATCTATTTTTAAAGAAAGCAAAATATATAACGATCCAGTGT	1653
Db	1381	CGTCATTTATACAAAATTAATCATCTATTTTTAAAGAAAGCAAAATATATAACGATCCAGTGT	1440
Qy	1654	GGGCTCGCTGCCCAAGTATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTAACCACT	1713
Db	1441	GGGCTCGCTGCCCAAGTATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTAACCACT	1500
Qy	1714	GGTGAATGGGAAGTCTTTGGCCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG	1773
Db	1501	GGTGAATGGGAAGTCTTTGGCCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG	1560
Qy	1774	CTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATACCTGTAGTAGTCACT	1833
Db	1561	CTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATACCTGTAGTAGTCACT	1620
Qy	1834	TACGTAATCTCGGAGGAGTGAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1893
Db	1621	TACGTAATCTCGGAGGAGTGAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1680
Qy	1894	ATCAGTCAAGCTGTGACTTTCTTTAAGTAAGTATAGTAAACCAAGAAATCCAACACTGT	1953
Db	1681	ATCAGTCAAGCTGTGACTTTCTTTAAGTAAGTATAGTAAACCAAGAAATCCAACACTGT	1740
Qy	1954	GTGTCCTTTTCAAGCTATCAAGCTCTGAAGATGACCAACTTGCAAAAACAAAGAAATTT	2013
Db	1741	GTGTCCTTTTCAAGCTATCAAGCTCTGAAGATGACCAACTTGCAAAAACAAAGAAATTT	1800
Qy	2014	TGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2073
Db	1801	TGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	1860
Qy	2074	TCCTTTTGAAGTACTGCTGATTTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	2133
Db	1861	TCCTTTTGAAGTACTGCTGATTTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	1920
Qy	2134	CAGCCTGGAAGAAATATCTACTGCTGTTTTCATATATGGTGGTCTCAGGTGCTGTTG	2193
Db	1921	CAGCCTGGAAGAAATATCTACTGCTGTTTTCATATATGGTGGTCTCAGGTGCTGTTG	1980
Qy	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCCGCTTTGAATACCTTAGCCTCTAGGT	2253
Db	1981	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCCGCTTTGAATACCTTAGCCTCTAGGT	2040
Qy	2254	TATGCTGTTGTAGTATAGACAAACAGGGGATCCTGTACCCGAGGCTTTAAATTTGAAGGC	2313
Db	2041	TATGCTGTTGTAGTATAGACAAACAGGGGATCCTGTACCCGAGGCTTTAAATTTGAAGGC	2100
Qy	2314	GCCTTTTAAATATATAATGGGTCAAATAGAAATTCACGATCAGGTGGAGGATCCCAATAT	2373
Db	2101	GCCTTTTAAATATATAATGGGTCAAATAGAAATTCACGATCAGGTGGAGGATCCCAATAT	2160
Qy	2374	CTAGCTTCTCGATATGATTTCAATTCATTCAGTATGCTGGGCATCCACGGCTGCTCTAT	2433
Db	2161	CTAGCTTCTCGATATGATTTCAATTCATTCAGTATGCTGGGCATCCACGGCTGCTCTAT	2220
Qy	2434	GGAGGATACCTCTCCCTGATGCGCATTAATGCAGAGTTCAGATATCTTCAGGGTTGCTATT	2493
Db	2221	GGAGGATACCTCTCCCTGATGCGCATTAATGCAGAGTTCAGATATCTTCAGGGTTGCTATT	2280
Qy	2494	GCTGGGGCCCCAGTCTCTGAGGATCTTCTATGATACAGGATACACGGAACGTTTATATG	2553
Db	2281	GCTGGGGCCCCAGTCTCTGAGGATCTTCTATGATACAGGATACACGGAACGTTTATATG	2340
Qy	2554	GGTCACCTCTGACCAAGTGAACAGGGCTATTAATCTAGGATCTGTGGCCATGCAAGCAGAA	2613
Db	2341	GGTCACCTCTGACCAAGTGAACAGGGCTATTAATCTAGGATCTGTGGCCATGCAAGCAGAA	2400
Qy	2614	AAGTTCCTCTGACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2673
Db	2401	AAGTTCCTCTGACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2460

QY 2674 CATTTTGCACATACCACTATATTAATCTAGTCTTTTGTAGTGGGCTGGAAAGCCATATGAT 2733
DB 2461 CATTTTGCACATACCACTATATTAATCTAGTCTTTTGTAGTGGGCTGGAAAGCCATATGAT 2520
QY 2734 TTACAGATCTATCTCTCAGGAGAGACACAGCATAAGAGTTCTTGAATCGGGAGAACATTTAT 2793
DB 2521 TTACAGATCTATCTCTCAGGAGAGACACAGCATAAGAGTTCTTGAATCGGGAGAACATTTAT 2590
QY 2794 GAACCTGCATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGATTATGCTGCTCTAAAA 2853
DB 2581 GAACCTGCATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGATTATGCTGCTCTAAAA 2640
QY 2854 GTGATATAA 2862
DB 2641 GTGATATAA 2649

RESULT 12
ABK83322
ID ABK83322 standard; cDNA; 2671 BP.
XX
AC ABK83322;
DT 12-AUG-2002 (first entry)
XX
DE cDNA encoding human DPPIV related serine protease DPPR-1.
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPPR;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX Homo sapiens.
OS
XX WO200231134-A2.
PN 18-APR-2002.
XX
PD 12-OCT-2001; 2001WO-US031874.
PF
XX 12-OCT-2000; 2000US-0240117P.
PR
XX (FERR) FERRING BV.
PA
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
PI
XX WPI; 2002-444178/47.
DR P-PSDB; ABG61591.
XX
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
XX Claim 1; Page 53-54; 113pp; English.
XX
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPR). The dipeptidyl peptidase IV-related proteins (DPPR) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPPR proteins

XX SQ Sequence 2671 BP; 805 A; 524 C; 594 G; 748 T; 0 U; 0 Other;
Query Match 84.8%; Score 2644.4; DB 6; Length 2671;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 209 GCAACATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAACTGCCG 268
DB 3 GTACCATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAACTGCCG 62
QY 269 ACTGTGAGGAGAATATTGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATGTTGAGC 328
DB 63 ACTGTGAGGAGAATATTGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATGTTGAGC 122
QY 329 GGTATTCTCTGGAGTCAGCTTAAAAAGCTGTTGCCGATACAGAAATATCATGGCTACA 388
DB 123 GGTATTCTCTGGAGTCAGCTTAAAAAGCTGTTGCCGATACAGAAATATCATGGCTACA 182
QY 389 TGATGGCTAAGGCACCAATGATTTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTC 448
DB 183 TGATGGCTAAGGCACCAATGATTTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTC 242
QY 449 ATTCTGACAGAAATCTATTACCTTCCCATGTCGTGAGAAACAGAGAAATATACACTGTTT 508
DB 243 ATTCTGACAGAAATCTATTACCTTCCCATGTCGTGAGAAACAGAGAAATATACACTGTTT 302
QY 509 ATTCTGAAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGAAGCCTC 568
DB 303 ATTCTGAAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGAAGCCTC 362
QY 569 TTTTGGATCTTTTTCAGGCAACACTGCACTATGGAATGATTTCTCGAGAGAGAACTAT 628
DB 363 TTTTGGATCTTTTTCAGGCAACACTGCACTATGGAATGATTTCTCGAGAGAGAACTAT 422
QY 629 TAAGGAAAGAAACCGCAATGGAACAGTCGGAATTTGCTTACGATATATCACCAAGAA 688
DB 423 TAAGGAAAGAAACCGCAATGGAACAGTCGGAATTTGCTTACGATATATCACCAAGAA 482
QY 689 GTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTAAGAGATGGAGGCGCAC 748
DB 483 GTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTAAGAGATGGAGGCGCAC 542
QY 749 AAGGATTTTACGCAACCACTTTTAAGGCGCCAACTTAGTGGAAACTAGTGTCCCAACATAC 808
DB 543 AAGGATTTTACGCAACCACTTTTAAGGCGCCAACTTAGTGGAAACTAGTGTCCCAACATAC 602
QY 809 GGATGGATCCAAATTTATGCCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACG 868
DB 603 GGATGGATCCAAATTTATGCCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACG 662
QY 869 ATATTGGATATCTAACTATGTAACGTAAGAGAAAGAGAGACTCACCTTATGTGCACAATG 928
DB 663 ATATTGGATATCTAACTATGTAACGTAAGAGAAAGAGAGACTCACCTTATGTGCACAATG 722
QY 929 AGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTGGCTTGTCTTCTCCAAG 988
DB 723 AGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTGGCTTGTCTTCTCCAAG 782
QY 989 AGAATTTGATAGATATTTCTGGCTATTTGCTGCTCCAAAGCTGAACACTCCCAAGT 1048
DB 783 AGAATTTGATAGATATTTCTGGCTATTTGCTGCTCCAAAGCTGAACACTCCCAAGT 842
QY 1049 GTGGTAAAAATCTTAGAATTTTATATGAAGAAATATGATCTAGCTGGAGTAATTC 1108
DB 843 GTGGTAAAAATCTTAGAATTTTATATGAAGAAATATGATCTAGCTGGAGTAATTC 902
QY 1109 ATGTTACATCCCTATGTTGGAAACAAGAGGCGCAATTCATTCGTTTATCTTAAACAG 1168
DB 903 ATGTTACATCCCTATGTTGGAAACAAGAGGCGCAATTCATTCGTTTATCTTAAACAG 962
QY 1169 GTACAGCAATCTCTAAAGTCACCTTTTAAAGATGTGCAAAATATGATGTGCTGAGGAA 1228

XX	22-JAN-2002; 2002GB-00001404.	QY	634	GAAGAAACGCATTGGAAACAGTCGGAATTCCTTACGATTATCAACCAAGGAAGTGA	693
PF		DB	421	GAAGAAACGCATTGGAAACAGTCGGAATTCCTTACGATTATCAACCAAGGAAGTGA	480
XX	23-JAN-2001; 2001GB-00001760.	QY	694	ACATTTCTGTTTCAAGCCGGTAGTGAATTTATACGTTAAAGATGAGGAGGCCACAAGGA	753
PR	(GLAX) GLAXO GROUP LTD.	DB	481	ACATTTCTGTTTCAAGCCGGTAGTGAATTTATACGTTAAAGATGAGGAGGCCACAAGGA	540
XX	Edbrooke MR, Lewis AP;	QY	754	TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCAACATACGGATG	813
PI		DB	541	TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCAACATACGGATG	600
XX	WPI: 2003-150703/15.	QY	814	GATCCAAAATATGCCCCGCTGATCCAGACTGGATTCCTTTTATACATAGAACGATATT	873
DR	P-PSDB; ASU07720.	DB	601	GATCCAAAATATGCCCCGCTGATCCAGACTGGATTCCTTTTATACATAGAACGATATT	660
XX	Identifying modulators of serine protease activity useful for treating	QY	874	TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCATTTATGTGCAATAGAGCTA	933
PT	musculoskeletal diseases, by contacting cell expressing a novel serine	DB	661	TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCATTTATGTGCAATAGAGCTA	720
PT	protease polypeptide with a compound and monitoring serine protease	QY	934	GCCAAACATGGAAGAAAGATGCGAGATCAGCTGGAGTCGCTACCTTTTGTTCCTCAAGAGAA	993
PS	Claim 12; Page 22-26; 38pp; English.	DB	721	GCCAAACATGGAAGAAAGATGCGAGATCAGCTGGAGTCGCTACCTTTTGTTCCTCAAGAGAA	780
XX	The invention relates to a method of identifying a substance that	QY	994	TTTGATAGATATCTCGCTATTTGGTGTGTCCTCAAGAGCTGAAACAACTCCCAAGTCGGT	1053
CC	modulates serine protease activity, comprising contacting a cell, such as	DB	781	TTTGATAGATATCTCGCTATTTGGTGTGTCCTCAAGAGCTGAAACAACTCCCAAGTCGGT	840
CC	a neuronal cell, lung cell, intestinal cell or a cell infected with a	QY	1054	AAAAATCTTACAAATCTTATATGAAGAAATGATGAATCTGAGGTGGAATATTATCATGTT	1113
CC	virus, expressing a serine protease polypeptide (H1PHUM 46), or its	DB	841	AAAAATCTTACAAATCTTATATGAAGAAATGATGAATCTGAGGTGGAATATTATCATGTT	900
CC	variant having dipeptidyl peptidase activity, or a serine protease	QY	1114	ACATCCCTATGTTGGAAACAGAGGCGAGATTCATTCGTTATCTTAAACACAGGTACA	1173
CC	isolated from the cell with a test substance and monitoring for serine	DB	901	ACATCCCTATGTTGGAAACAGAGGCGCGAGATTCATTCGTTATCTTAAACACAGGTACA	960
CC	protease activity. The method is useful for identifying a substance that	QY	1174	GCAATCTCTAAAGTCATCTTTTAAGATCTCGAATATGATGATGATGATGATGATGATG	1233
CC	modulates serine protease activity. A modulator of the serine protease is	DB	961	GCAATCTCTAAAGTCATCTTTTAAGATCTCGAATATGATGATGATGATGATGATGATG	1020
CC	useful in the manufacture of a medicament for treatment or prophylaxis of	QY	1234	ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGGAGATTCCTTATTTGAAGGTTGAA	1293
CC	a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus	DB	1021	ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGGAGATTCCTTATTTGAAGGTTGAA	1080
CC	infection, Alzheimer's disease, paraspinal nuclear palsy, myotonic	QY	1294	TATATTGCGAGAGCTGGATGGACTCCTGAGGGAATAATATGCTTGGTCCATCTCTACTAGAT	1353
CC	dystrophy, Huntington's disease or amyotrophic lateral sclerosis.	DB	1081	TATATTGCGAGAGCTGGATGGACTCCTGAGGGAATAATATGCTTGGTCCATCTCTACTAGAT	1140
CC	Additional disease that may be treated using modulators of the serine	QY	1354	CGCTCCAGACTCGCTCCTACAGATGATGTTGATCTCACCTGAAATTTATTTATCCAGTAGAA	1413
CC	protease include malabsorption syndromes, irritable bowel syndrome, lung	DB	1141	CGCTCCAGACTCGCTCCTACAGATGATGTTGATCTCACCTGAAATTTATTTATCCAGTAGAA	1200
CC	disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,	QY	1414	GATGATGTTTATGGAAGGCGAGACTCAATGAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG	1473
CC	rectal polyps, small bowel tumours, colorectal tumours, anaemia,	DB	1201	GATGATGTTTATGGAAGGCGAGACTCAATGAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG	1260
CC	dyslexia, teroid lipofuscinosis, allergic encephalomyelitis, and multiple	QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTATGTTT	1533
CC	scleriosis. The present sequence represents cDNA encoding the human serine	DB	1261	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTATGTTT	1320
CC	protease H1PHUM46 which is located on chromosome 15q21-q22	QY	1534	CCCCAAGTCAGAGAGGAAATTCAGTTTATTTTGGCTCTGATTCCTGAAATCAAGGTTTC	1593
XX	Sequence 2649 BP; 803 A; 514 C; 585 G; 747 T; 0 U; 0 Other;	DB	1321	CCCCAAGTCAGAGAGGAAATTCAGTTTATTTTGGCTCTGATTCCTGAAATCAAGGTTTC	1380
XX	Query Match 84.7%; Score 2644.2; DB 8; Length 2649;	QY	1594	CGTCATTTTATCAAAATTCATCTATTTTAAAGGAAACCAATATTAACCATCCAGTGGT	1653
XX	Best Local Similarity 99.9%; Pred. No. 0;	DB	1381	CGTCATTTTATCAAAATTCATCTATTTTAAAGGAAACCAATATTAACCATCCAGTGGT	1440
XX	Matches 2646; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	QY	1654	GGGCTGCTGCTCCAAAGTCATTTCAAGTGTCTTATCAAGAGGAGTAGCAATTTACCAGT	1713
QY	214 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTTGAGATATTTGAACTGGGACTGT	DB	1441	GGGCTGCTGCTCCAAAGTCATTTCAAGTGTCTTATCAAGAGGAGTAGCAATTTACCAGT	1500
DB	1 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTTGAGATATTTGAACTGGGACTGT	QY	1714	GGTGAATGGGAAGTCTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG	1773
QY	274 GAGGAGATATTTGAATCAAGATCGGCTAAATTTGAGGCTTTTATGTTGAGCGGTAT				
DB	61 GAGGAGATATTTGAATCAAGATCGGCTAAATTTGAGGCTTTTATGTTGAGCGGTAT				
QY	334 TCCTGGAGTCAGCTTAAAGGCTGCTTCCGATACCGAAATAATATCATGCTACATGATG				
DB	121 TCCTGGAGTCAGCTTAAAGGCTGCTTCCGATACCGAAATAATATCATGCTACATGATG				
QY	394 GCTAAGGCAACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTC				
DB	181 GCTAAGGCAACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTC				
QY	454 GAGAGATCTATTACCTTCCCATGCTGCTGAGAACAGAGAAATAACACTGTTTATTTCT				
DB	241 GAGAGATCTATTACCTTCCCATGCTGCTGAGAACAGAGAAATAACACTGTTTATTTCT				
QY	514 GAAATTCGAAACTATCATAGCAGCAGCTTTAATGCTCTCTTGGAGGCTCTTTTGG				
DB	301 GAAATTCGAAACTATCATAGCAGCAGCTTTAATGCTCTCTTGGAGGCTCTTTTGG				
QY	574 GATCTTTTTCAGGCAACATCGGACTGATGATGATTTCTCCGAGAGGAAGCAATTTAAGA				
DB	361 GATCTTTTTCAGGCAACATCGGACTGATGATGATTTCTCCGAGAGGAAGCAATTTAAGA				

Db 1501 GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAAATATCAAAGTTGATGAAGTCAGAAGG 1560
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACCTAGTCAGT 1833
Db 1561 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACCTAGTCAGT 1620
QY 1834 TACGTAATCTCGAGAGGTGACCAAGCTGACTGACCGTGGGTATCTCAATTTCTTGCTGC 1893
Db 1621 TACGTAATCTCGAGAGGTGACCAAGCTGACTGACCGTGGGTATCTCAATTTCTTGCTGC 1680
QY 1894 ATCAGTCAGCAGCTGCTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGT 1953
Db 1681 ATCAGTCAGCAGCTGCTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGT 1740
QY 1954 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGGAATTT 2013
Db 1741 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGGAATTT 1800
QY 2014 TGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTTGACTATATCTCTCCAGAAATTTTC 2073
Db 1801 TGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTTGACTATATCTCTCCAGAAATTTTC 1860
QY 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGATTTGGATGCTCTACAGCCTCATGATCTA 2133
Db 1861 TCTTTTGAAGTACTACTGGATTTTACATTTGATTTGGATGCTCTACAGCCTCATGATCTA 1920
QY 2134 CAGCTCGAAGAAATATCTTACCTGCTGCTGTATATATGCTGCTCTCAGGTGCAAGTTC 2193
Db 1921 CAGCTCGAAGAAATATCTTACCTGCTGCTGTATATATGCTGCTCTCAGGTGCAAGTTC 1980
QY 2194 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATATCCCTAGCCTCTCTAGT 2253
Db 1981 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATATCCCTAGCCTCTCTAGT 2040
QY 2254 TATGTGGTTAGTAGATAGACCAAGGGGATCTGTGTCACCGAGGGCTTAAATTTGAAGGC 2313
Db 2041 TATGTGGTTAGTAGATAGACCAAGGGGATCTGTGTCACCGAGGGCTTAAATTTGAAGGC 2100
QY 2314 GCCTTTAATATTAATGGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAATAT 2373
Db 2101 GCCTTTAATATTAATGGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAATAT 2160
QY 2374 CTAGCTCTCGATATGATTTCAATTTGACTTAGATCTGTGGGCATCCACGGCTGCTCTAT 2433
Db 2161 CTAGCTCTCGATATGATTTCAATTTGACTTAGATCTGTGGGCATCCACGGCTGCTCTAT 2220
QY 2434 GGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTFCAGATATCTTCAGGGTTGCTATT 2493
Db 2221 GGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTFCAGATATCTTCAGGGTTGCTATT 2280
QY 2494 GCTGGGGCCCGTCACTCTGTGATCTTCTATGATACAGATACACGGNAGCTTATATG 2553
Db 2281 GCTGGGGCCCGTCACTCTGTGATCTTCTATGATACAGATACACGGNAGCTTATATG 2340
QY 2554 GGTCACTCTGACCAAGTAAACAGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA 2613
Db 2341 GGTCACTCTGACCAAGTAAACAGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA 2400
QY 2614 AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAATGTC 2673
Db 2401 AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAATGTC 2460
QY 2674 CATTTTGACATACAGATATATTAATGATGTTTTTGTAGGGCTGGAAGCCATATGAT 2733
Db 2461 CATTTTGACATACAGATATATTAATGATGTTTTTGTAGGGCTGGAAGCCATATGAT 2520
QY 2734 TTACAGATCTCTCAGGAGACACAGATAGAGTCTCTCAATCGGAGAACATTAT 2793
Db 2521 TTACAGATCTCTCAGGAGACACAGATAGAGTCTCTCAATCGGAGAACATTAT 2580
QY 2794 GAACTGCATCTTTTGCATCTTACCTTCAAGAAACCTTGGATCAGGTATGCTGCTCTAAAA 2853

Db 2581 GAACCTGCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGCTATTGCTCTAAAA 2640
QY 2854 GTGATATAA 2862
Db 2641 GTGATATAA 2649
RESULT 14
AAH99935
ID AAH99935 standard; cDNA; 2643 BP.
XX AC AAH99935;
XX AC AAH99935;
DT 12-APR-2002 (first entry)
XX
Coding sequence of 21953 human prollyl oligopeptidase.
XX 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
XX cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease; ss.
XX Homo sapiens.
XX WO200179473-A2.
XX 25-OCT-2001.
XX 11-APR-2001; 2001WO-US040483.
XX 18-APR-2000; 2000US-0197508P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Meyers RA, Williamson M;
XX WPI; 2002-034353/04.
XX P-PSDB; AAG78415.
XX New polypeptides 21953, member of human prollyl oligopeptidase family,
XX useful as diagnostic targets and therapeutic agents for controlling
XX cancer, lymphoma and leukemia.
XX Claim 2; Page; 121pp; English.
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.
XX Which is cytosolic, antidiabetic, antiarthritic, neuroprotective,
XX antithyroid, dermatological, antipsoriatic, antiasthmatic,
XX ophthalmological, antiinflammatory, neurotropic, antiparkinsonian,
XX anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,
XX antiatherosclerotic, anorectic and metabolic in its action. Uses include
XX gene therapy, expression or activity of 21953 protein modulator, it is
XX useful for identifying a compound which binds to it and can be used in
XX preventing, treating or detecting a cellular proliferative or
XX differentiative disorder. The 21953 molecules can act as novel diagnostic
XX targets and therapeutic agents for controlling disorders associated with
XX the aberrant activity or degradation of peptide hormones e.g., disorders
XX associated with cell differentiation and proliferation such as cancer,
XX immune function, reproductive, neurological and cardiovascular function.
XX The 21953 molecules are thus useful for treating and preventing cellular
XX proliferative and differentiative disorders, haematopoietic neoplastic
XX disorders, immune disorders such as autoimmune diseases, diabetes
XX mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
XX neuronal disorders, demyelinating diseases, vascular disorders and
XX metabolism or pain disorders. This sequence represents the cDNA encoding
XX sequence of 21953 human prollyl oligopeptidase. This sequence represents
XX the coding sequence of 21953, being the sequence in between the start and
XX the stop codon of the sequence represented in AAH99934
XX Sequence 2643 BP; 800 A; 514 C; 585 G; 744 T; 0 U; 0 Other;

Query Match										
Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;										
Matches 2641; Conservative 0; Mismatches 2; Indels 0; Gaps 0;										
Qy	217	GCAGCGCAATGGAACAGAACAGCTGGGTGGATATTTGAACTGCGGACTGTGAG	276	Db	1021	GATGTCATAGATAAGGAAC	TAATTC	CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATAT	1080	
Db	1	GCAGCGCAATGGAACAGAACAGCTGGGTGGATATTTGAACTGCGGACTGTGAG	60	Qy	1297	ATTGCCAGAGCTGGATGGAC	TCTCTGAGGAAAAATATGCTTTGGTCCATCTACTAGATCGC	1356		
Qy	277	GAGATATTTGAATCAGAGATCGGCTTAATTTGGAGCCCTTTTATGTTGAGCGGTATCC	336	Db	1081	ATTGCCAGAGCTGGATGGAC	TCTCTGAGGAAAAATATGCTTTGGTCCATCTACTAGATCGC	1140		
Db	61	GAGATATTTGAATCAGAGATCGGCTTAATTTGGAGCCCTTTTATGTTGAGCGGTATCC	120	Qy	1357	TCCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAAGAT	1416			
Qy	337	TGGAGTCAGCTTAAAGAGCTGCTCCGATACAGAAATATCATGAGTACATGATGGCT	396	Db	1141	TCCCAGAGCTCGCCTGCAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAAGAT	1200			
Db	121	TGGAGTCAGCTTAAAGAGCTGCTCCGATACAGAAATATCATGAGTACATGATGGCT	180	Qy	1417	GATGTTATGGAAAGGAGAGACTCAATTGAGTCAGTGCCTGATCTGTGAGCGCCACTAAAT	1476			
Qy	397	AAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGAGCTCATTCAGAC	456	Db	1201	GATGTTATGGAAAGGAGAGACTCAATTGAGTCAGTGCCTGATCTGTGAGCGCCACTAAAT	1260			
Db	181	AAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGAGCTCATTCAGAC	240	Qy	1477	ATCTATGAAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTTCCC	1536			
Qy	457	AGAAATCTATTACCTTGGCATGTCTGGTGAAGAACAGAGAAAAATACACTGTTTTATCTGAA	516	Db	1261	ATCTATGAAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTTCCC	1320			
Db	241	AGAAATCTATTACCTTGGCATGTCTGGTGAAGAACAGAGAAAAATACACTGTTTTATCTGAA	300	Qy	1537	CAAGTCAGGAGAGGAAATTTGAGTTTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGT	1596			
Qy	517	ATTTCCAAATCTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGAT	576	Db	1321	CAAGTCAGGAGAGGAAATTTGAGTTTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGT	1380			
Db	301	ATTTCCAAATCTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGAT	360	Qy	1597	CAATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGATCCAGTGGTGG	1656			
Qy	577	CTTTTTCAGGCAACACTGGACTATGGAATGTAATCTCGAGAGAGAAAGTAATTAAGAGAA	636	Db	1381	CAATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGATCCAGTGGTGG	1440			
Db	361	CTTTTTCAGGCAACACTGGACTATGGAATGTAATCTCGAGAGAGAAAGTAATTAAGAGAA	420	Qy	1657	CTGCCTGCTCCAAGTGATTTCAAAGTGCTCTATCAAGAGGAGATAGCAATTTACAGTGGT	1716			
Qy	637	AGAAACCGATTTGGAACAGTCGGAATGCTTCTTACGATTAATCACCAGGAGTGGAAACA	696	Db	1441	CTGCCTGCTCCAAGTGATTTCAAAGTGCTCTATCAAGAGGAGATAGCAATTTACAGTGGT	1500			
Db	421	AGAAACCGATTTGGAACAGTCGGAATGCTTCTTACGATTAATCACCAGGAGTGGAAACA	480	Qy	1717	GAATGGAGATTTCTTGGCGGCGGATGATTAATATCCAAAGTTGATGAAATCAGAAGGCTG	1776			
Qy	697	TTTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCGCAAGGATTT	756	Db	1501	GAATGGAGATTTCTTGGCGGCGGATGATTAATATCCAAAGTTGATGAAATCAGAAGGCTG	1560			
Db	481	TTTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCGCAAGGATTT	540	Qy	1777	GTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAAGTTAC	1836			
Qy	757	AGCACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCTCCAACTACGGATGGAT	816	Db	1561	GTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAAGTTAC	1620			
Db	541	AGCACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCTCCAACTACGGATGGAT	600	Qy	1837	GTAAATCTCTGGAGAGGTCACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGCATC	1896			
Qy	817	CCAAAATTTATGCCCGCTGTATCCAGACTGGATTGCTTTTATACATAGCAACGATATTTGG	876	Db	1621	GTAAATCTCTGGAGAGGTCACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGCATC	1680			
Db	601	CCAAAATTTATGCCCGCTGTATCCAGACTGGATTGCTTTTATACATAGCAACGATATTTGG	660	Qy	1897	AGTCAGCATGTGACTCTCTTTTAAAGTAAGTATAGTAACCCAGAGAGATCCACACTGTGTG	1956			
Qy	877	ATATCTAACATCGTAACAGAGAGAAAGGAGACTCACATTTATGTGCACAATGAGCTAGCC	936	Db	1681	AGTCAGCATGTGACTCTCTTTTAAAGTAAGTATAGTAACCCAGAGAGATCCACACTGTGTG	1740			
Db	661	ATATCTAACATCGTAACAGAGAGAAAGGAGACTCACATTTATGTGCACAATGAGCTAGCC	720	Qy	1957	TCCCTTTACAAGCTATCAAGTCTCGAAGATGACCCACTTTCGCAAAAACAAAGGAAATTTGG	2016			
Qy	937	AACATGGAAGAGATGCGCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAAGAAATTT	996	Db	1741	TCCCTTTACAAGCTATCAAGTCTCGAAGATGACCCACTTTCGCAAAAACAAAGGAAATTTGG	1800			
Db	721	AACATGGAAGAGATGCGCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAAGAAATTT	780	Qy	2017	GCCACATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCAGAAAATTTTCTCT	2076			
Qy	997	GATAGATTTCTGGCTATGCTGGTGTCCAAAGCTGAAACACTCCAGTGGTGTAAA	1056	Db	1801	GCCACATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCAGAAAATTTTCTCT	1860			
Db	781	GATAGATTTCTGGCTATGCTGGTGTCCAAAGCTGAAACACTCCAGTGGTGTAAA	840	Qy	2077	TTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAG	2136			
Qy	1057	ATTTCTAGATTTCTATATGAAGAAATGATGATCTGAGTGGAAATTTATCTGTTTACA	1116	Db	1861	TTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAG	1920			
Db	841	ATTTCTAGATTTCTATATGAAGAAATGATGAAATCTGAGTGGAAATTTATCTGTTTACA	900	Qy	2137	CCTGGAAGCAAAATATCTACTGTGCTGTTTCATATATGTTGGTGTCTCTCAGGTGTCAGTGGTG	2196			
Qy	1117	TCCCTATGTTGAAACCAAGAGGCGCAGATTCATTCGGTTATCCTAAAACAGGTACAGCA	1176	Db	1921	CCTGGAAGCAAAATATCTACTGTGCTGTTTCATATATGTTGGTGTCTCTCAGGTGTCAGTGGTG	1980			
Db	901	TCCCTATGTTGAAACCAAGAGGCGCAGATTCATTCGGTTATCCTAAAACAGGTACAGCA	960	Qy	2197	AATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCTAGGCTCTCTAGGTTAT	2256			
Qy	1177	AATCCTAAAGTCACCTTTTAAAGATGTCAAGAAATATGATGCTGAAAGGAGGATCATAT	1236	Db	1981	AATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCTAGGCTCTCTAGGTTAT	2040			
Db	961	AATCCTAAAGTCACCTTTTAAAGATGTCAAGAAATATGATGCTGAAAGGAGGATCATAT	1020	Qy	2257	GTGGTTGTAGTAGACAAAGGGGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCC	2316			
Qy	1237	GATGTCATAGATAAGGAAC	TAATTC	Db	2041	GTGGTTGTAGTAGACAAAGGGGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCC	2100			
				Qy	2317	TTTAAATATAAATGGGTC	CAAAATG	ACGAT	CAGGTGGAGGAGCTCCATATCTA	2376

Db	2101	TTTAAATATATAAATGGGTCAATAGAAATGACGATCAGGTGGAGGACTCAATATCTA	2160
Qy	2377	GCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGA	2436
Db	2161	GCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGA	2220
Qy	2437	GGATACCTCTCCCTGATGGCATTAATGACGAGGTGAGATATCTTCAGGGTTGCTATTGCT	2496
Db	2221	GGATACCTCTCCCTGATGGCATTAATGACGAGGTGAGATATCTTCAGGGTTGCTATTGCT	2280
Qy	2497	GGGGCCCCAGTCACTCTGTGGATCTTATGATACAGGATACACGGAACGTTATATGGGT	2556
Db	2281	GGGGCCCCAGTCACTCTGTGGATCTTATGATACAGGATACACGGAACGTTATATGGGT	2340
Qy	2557	CACCTGCACAGAAATGAACAGGCTATTAATTAGGATCTGTGGCCATGCAAGCAGAAAAG	2616
Db	2341	CACCTGCACAGAAATGAACAGGCTATTAATTAGGATCTGTGGCCATGCAAGCAGAAAAG	2400
Qy	2617	TTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGCCAT	2676
Db	2401	TTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGCCAT	2460
Qy	2677	TTTGCATACACAGTATATTAATGAGTCTTTTATGAGGGCTGGAAGCCATATGATTTA	2736
Db	2461	TTTGCATACACAGTATATTAATGAGTCTTTTATGAGGGCTGGAAGCCATATGATTTA	2520
Qy	2737	CAGATCTATCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAAATATATGAA	2796
Db	2521	CAGATCTATCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAAATATATGAA	2580
Qy	2797	CTGCATCTTTTGCATACCTTTCAAGAAACCTTTGGATCAGTATTGCTGCTCTAAAAGTG	2856
Db	2581	CTGCATCTTTTGCATACCTTTCAAGAAACCTTTGGATCAGTATTGCTGCTCTAAAAGTG	2640
Qy	2857	ATA 2859	
Db	2641	ATA 2643	
RESULT 15			
ID	ABK83325		
XX	ABK83325 standard; cDNA; 4523 BP.		
AC	ABK83325;		
DT	12-AUG-2002 (first entry)		
XX	cDNA encoding human DPPP-1 splice variant #1.		
DE	Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;		
KW	DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;		
KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;		
KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;		
KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;		
KW	dyskinesia; reproductive disorder; inflammatory disorder;		
KW	metabolic disorder; gene; ss.		
OS	Homo sapiens.		
XX	WO200231134-A2.		
PN	18-APR-2002.		
PD	12-OCT-2001; 2001WO-US031874.		
PF	12-OCT-2000; 2000US-0240117P.		
PR	(FERR) FERRING BV.		
XX	Qi S, Akinsanya KO, Riviere PJ, Junien J;		
XX	WPI: 2002-444178/47.		
PI	P-PSDB; ABG61594.		
DR			

New dipeptidyl peptidase IV-related proteins and nucleic acids encoding the proteins, useful for treating e.g. fungal, bacterial, protozoan and viral infections, cancers, allergies, neurological disorders, or pain.

Disclosure; Page 61-62; 113pp; English.

The present invention relates to the isolation of novel human serine proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic acids encoding them are useful for treating infections such as fungal, bacterial, protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes, precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, stroke, ulcers, asthma, allergies, cancers, migraine, vomiting, psychotic and neurological disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias. These may also be used in discovering therapeutic agents for the treatment of reproductive, inflammatory and metabolic disorders. ABK83322 -ABK83343 encode human DPPP proteins

CC Sequence 4523 BP; 1384 A; 828 C; 940 G; 1371 T; 0 U; 0 Other;

Query Match 80.1%; Score 2500.4; DB 6; Length 4523;

Best Local Similarity 90.2%; Pred. No. 0;

Matches 2815; Conservative 0; Mismatches 1; Indels 304; Gaps 1;

Qy 1 AAGTGCTAAAGCTCCGAGGCAAGCGCTGCTACTGCGCGCGCTCTTTAGTGGCG 60

Db 1 AAGTGCTAAAGCTCCGAGGCAAGCGCTGCTACTGCGCGCGCTCTTTAGTGGCG 60

Qy 61 CGTTGCGCGCTGGGTGTTTCAACGCGCGCGCGCGCGAGGAGCACTGCAACGAGACCG 120

Db 61 CGTTGCGCGCTGGGTGTTTCAACGCGCGCGCGCGAGGAGCACTGCAACGAGACCG 120

Qy 121 GAGTGAGGCGCGCGAGCATGAGAGCGCGCGCGCGCGCTGCTGAGCGAGCGTGGGACCG 180

Db 121 GAGTGAGGCGCGCGAGCATGAGAGCGCGCGCGCGCTGCTGAGCGAGCGTGGGACCG 180

Qy 181 TCGGCGCGCGCGCGGAGGAGAAATGCAACATGCGAGCAATGGAACAGAAACAGACAG 240

Db 181 TCGGCGCGCGCGCGGAGGAGAAATGCAACATGCGAGCAATGGAACAGAAACAGACAG 240

Qy 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAGAGATCGG 300

Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAGAGATCGG 300

Qy 301 CTTAAATTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTGCTT 360

Db 301 CTTAAATTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTGCTT 360

Qy 361 GCGGATACAGAAATATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Db 361 GCGGATACAGAAATATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Qy 421 GTGAAGAGGATGATCCAGATGAGCTCATTCAGACAGATCTATTACCTTGGCATGCTCT 480

Db 421 GTGAAGAGGATGATCCAGATGAGCTCATTCAGACAGATCTATTACCTTGGCATGCTCT 480

Qy 481 GGTGAGACAGAGAAATACACTGTTTATTTGTAATTCCTGAAATCTATCAATAGAGCA 540

Db 481 GGTGAGACAGAGAAATACACTGTTTATTTGTAATTCCTGAAATCTATCAATAGAGCA 540

Qy 541 GCAGTCTTAATGCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCAACCTGGATAT 600

Db 541 GCAGTCTTAATGCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCAACCTGGATAT 600

Qy 601 GGAATGATTTCTCGAGAGAGAACTATTTAAGAGAGAAAGAAACCGATTGGAACAGTCGGA 660

Db 601 GGAATGATTTCTCGAGAGAGAACTATTTAAGAGAGAAAGAAACCGATTGGAACAGTCGGA 660

Qy 661 ATTGCTTCTTACGATTTATCACCAGAGAGTGGAAATTTCTGTTTCAAGCGCGTAGTGGGA 720

Qy	2881	TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGGTTTAAATCAACAGAAAAACACAGA	2940
Db	2577	TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGGTTTAAATCAACAGAAAAACACAGA	2636
Qy	2941	ATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCCTCGAAAAATAAATGTGGTG	3000
Db	2637	ATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCCTCGAAAAATAAATGTGGTG	2696
Qy	3001	CCATGCAGGGGTCTACGGTTTGTGGTGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAA	3060
Db	2697	CCATGCAGGGGTCTACGGTTTGTGGTGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAA	2756
Qy	3061	TCAAATGATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTACTAAAAA	3120
Db	2757	TCAAATGATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTACTAAAAA	2816

Search completed: April 15, 2006, 00:48:48
Job time : 1779 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 19:46:05 ; Search time 10952 Seconds

(without alignments)
13328.679 Million cell updates/sec

Title: US-10-825-632-2

Perfect score: 3120

Sequence: 1 aadgtgtaagctccgagg.....agaattactaaaaaaa 3120

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 segs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_btc:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_est7:*
- 9: gb_gss1:*
- 10: gb_gss2:*
- 11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2446.8	78.4	2649	10	AY411615 Homo sapi
2	2381.4	76.3	2649	10	AY411616 Pan trogl
3	2251.8	72.2	3143	4	AK016546 Mus muscu
4	2158.6	69.2	5517	4	AK029788 Mus muscu
5	1985	63.6	2634	10	AY411617 Mus muscu
6	1948.8	62.5	2292	4	CR609512 full-leng
7	1217.4	39.0	1292	4	AF175225 Homo sapi
8	1200	38.5	1265	4	AF176779 Homo sapi
9	1132	36.3	1278	4	AF173382 Homo sapi
10	938.2	30.1	1042	3	BM557438 AGENCOURT
11	882.2	28.3	3327	4	AK050021 Mus muscu
12	880.6	28.2	3457	4	AK078301 Mus muscu
13	875.8	28.1	884	7	CR985650 CR985650
14	871.2	27.9	3376	4	AK050023 Mus muscu
15	848.6	27.2	957	5	BQ675006 AGENCOURT
16	848.2	27.2	881	7	CR990494 CR990494
17	845.6	27.1	910	5	BQ675260 AGENCOURT
18	844.4	27.1	910	5	BX372276 BX372276
19	837.4	26.8	985	3	BQ686650 AGENCOURT
20	827	26.5	864	1	AL542617 AL542617
21	818	26.2	1013	7	CN647232 ILLUMIGEN
22	818	26.2	1027	7	CN646612 ILLUMIGEN

c	23	812.2	26.0	910	2	BG479035	BG479035	602526111
	24	804.4	25.8	902	8	DN517062	DN517062	1257572 M
	25	803.6	25.8	914	5	BX390898	BX390898	EX390898
	26	801.8	25.7	835	7	CR998849	CR998849	CR998849
	27	801.2	25.7	1041	2	BI084090	BI084090	602869453
	28	784	25.1	968	5	BQ671635	BQ671635	AGENCOURT
c	29	775.4	24.9	855	2	BI084885	BI084885	602869453
	30	768	24.6	857	8	DN519771	DN519771	1262448 M
c	31	763.6	24.5	804	5	BG687687	BG687687	UI-CF-EC1
	32	761.6	24.4	789	2	BG709118	BG709118	602675382
	33	758.4	24.3	840	8	DN514280	DN514280	1251358 M
	34	756.6	24.3	820	7	CR994557	CR994557	CR994557
	35	750.4	24.1	753	1	AL040398	AL040398	DKP2p434A
	36	742.4	23.8	853	2	BI223892	BI223892	602941035
	37	736.4	23.6	742	7	CN427229	CN427229	170006002
c	38	734	23.5	735	1	AL043338	AL043338	DKF2p434O
	39	726.4	23.3	760	1	AI917735	AI917735	tt11d07.x
	40	719.6	23.1	826	3	BI553405	BI553405	603193351
	41	706.2	22.6	1002	2	BE733691	BE733691	601569288
	42	697.8	22.4	726	7	CN427231	CN427231	170004554
	43	694	22.2	724	7	CK000768	CK000768	AGENCOURT
	44	691.6	22.2	788	2	BG163397	BG163397	602338360
c	45	688.8	22.1	708	5	BU622228	BU622228	UI-H-PH1-

ALIGNMENTS

RESULT 1	AY411615	Homo sapiens DPP8 gene, genomic survey sequence.	2649 bp	DNA	linear	GSS 16-DEC-2003
AY411615	LOCUS	AY411615	2649 bp	DNA	linear	GSS 16-DEC-2003
DEFINITION	AY411615	Genomic survey sequence.				
ACCESSION	AY411615	GI:39767583				
VERSION	AY411615.1	GI:39767583				
KEYWORDS	GSS.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 2649)	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
AUTHORS	1 (bases 1 to 2649)	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios					
JOURNAL	Science 302 (5652), 1960-1963 (2003)					
PUBMED	14671302					
REFERENCE	2 (bases 1 to 2649)	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
AUTHORS	2 (bases 1 to 2649)	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission					
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA					
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.					
FEATURES	Location/Qualifiers					
source	1. .2649					
gene	/organism="Homo sapiens"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:9606"					
	<1. .>2649					
	/gene="DPP8"					
	/locus_tag="HCM4283"					
ORIGIN						
Query Match	78.4%	Score 2446.8	DB 10	Length 2649		
Best Local Similarity	92.4%	Pred. No. 0				
Matches 2448	Conservative 0	Mismatches 201	Indels 0	Gaps 0		

[illegible]

```
Db 2161 |||||CTAGCTTCGATATGATTTCAATGACCTAGATCGTGTGGGCATCCACGGCTGGTCCAT 2220
Qy 2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTTGAGGTTGCTATT 2493
Db 2221 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTTGAGGTTGCTATT 2280
Qy 2494 GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTTATG 2553
Db 2281 GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTTATG 2340
Qy 2554 GGTACCCCTGACAGAAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCGCAATGCAAGAA 2613
Db 2341 GGTACCCCTGACAGAAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCGCAATGCAAGAA 2400
Qy 2614 AAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGCTTTCTGGATGAGAATGTC 2673
Db 2401 NNNNNNNNNNNNACAAATCGTTTACTGCTCTTACATGCTTTCTGGATGAGAATGTC 2460
Qy 2674 CATTTTGCACATACCACTATATTAATGAGTGTGAGTGTGAGGCTGGAAAGCCATATGAT 2733
Db 2461 CATTTTGCACATACCACTATATTAATGAGTGTGAGTGTGAGGCTGGAAAGCCATATGAT 2520
Qy 2734 TTACAGATCTATCCTCAGAGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTAT 2793
Db 2521 TTACAGATCTATCCTCAGAGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTAT 2580
Qy 2794 GAATCGATCTTTTGCACCTACCTCAAGAAACCTTGCATCAGCTATGCTGCTCTAA 2853
Db 2581 GAATCGATCTTTTGCACCTACCTCAAGAAACCTTGCATCAGCTATGCTGCTCTAA 2640
Qy 2854 GTGATATA 2862
Db 2641 GTGATATA 2649
```

```
RESULT 2
AV411616 2649 bp DNA linear GSS 16-DEC-2003
LOCUS Pan troglodytes DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AV411616
VERSION AV411616.1 GI:39767584
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
REFERENCE 1 (bases 1 to 2649)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2649)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2649
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
```

```
gene <1..>2649
/genes="DPP8"
/locus_tag="HCM4283"
ORIGIN
Query Match 76.1%; Score 2381.4; DB 10; Length 2649;
Best Local Similarity 90.1%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 261; Indels 0; Gaps 0;
Qy 214 ATGCGACAGCAATGGAACACAGAACAGCTGGGTGTGGATATTTGAAACCTGCGACTGT 273
Db 1 ATGCGACAGCAATGGAACACAGAACAGCTGGGTGTGGATATTTGAAACCTGCGACTGT 60
Qy 274 GAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGAGGCTTTTATGTTGAGCGGTAT 333
Db 61 GAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGAGGCTTTTATGTTGAGCGGTAT 120
Qy 334 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCCAGAAATATCATGGCTACATGATG 393
Db 121 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCCAGAAATATCATGGCTACATGATG 180
Qy 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA 453
Db 181 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGGCTCATTTCA 240
Qy 454 GACAGAAATCTATTACCTTGCCATGCTCTGTGAGAACAGAGAAATATCATGTTTATTTCT 513
Db 241 GACAGAAATCTATTACCTTGNNNNNNNGGTGAGAACAGAGAAATATCATGTTTATTTCT 300
Qy 514 GAAATTCCTAAACATCAATAGACAGCAGCTTTAAATGCTCTCTTGGAGGCTCTTTTG 573
Db 301 GAAATTCCTAAACATCAATAGACAGCAGCTTTAAATGCTCTCTTGGAGGCTCTTTTG 360
Qy 574 GATCTTTTTCAGGCAACACCTGGACTATGAAATGTTTCTCGAGAGAGAAACTATTAAAGA 633
Db 361 GATCTTTTTCAGGCAACACCTGGACTATGAAATGTTTCTCGAGAGAGAAACTATTAAAGA 420
Qy 634 GAAAGAAACGCAATGGAACAGTCGGAATTTGCTTTTACGATTTATCACCAAGGAAGTCGA 693
Db 421 GAAAGAAACGCAATGGAACAGTCGGAATTTGCTTTTACGATTTATCACCAAGGAAGTCGA 480
Qy 694 ACATTTCTGTTTACGCGGTAGTGAATTTATCAGTAAAGATGGAGGGCCACAAAGA 753
Db 481 ACATTTCTGTTTCAAGCTGGTGTAGTGAATTTATCAGTAAAGATGGAGGGCCACAAAGA 540
Qy 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCCAACATACCGATG 813
Db 541 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCCAACATACCGATG 600
Qy 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 873
Db 601 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 660
Qy 874 TGGATATCTAACATCGTAAACAGAGAAAGAGAGCTCCTTATGTGACAAATGAGCTA 933
Db 661 TGGATATCTAACATCGTAAACAGAGAAAGAGAGACTCCTTATGTGACAAATGAGCTA 720
Qy 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGATCGCTACCTTTGTTCTCCCAAGAA 993
Db 721 GCCAACATGGAAGAGATGCCAGATCAGCTGGATTTCTACCTTTGTTCTCCCAAGAA 780
Qy 994 TTTGATAGATATTTCTGGCTATTGCTGGTGTGTCCTCAAGAGCTGAAACAACTCCCAAGTGGT 1053
Db 781 NNNNATAGATATTTCTGGCTATTGCTGGTGTGTCCTCAAGAGCTGAAACAACTCCCAAGTGGT 840
Qy 1054 AAAATTTCTAGAAATTTCTATATGAAGAAATATGATGAAATCTGAGGTGAAATTTATGTT 1113
Db 841 AAAATTTCTAGAAATTTCTATATGAAGAAATATGATGAAATCTGAGGTGAAATTTATGTT 900
Qy 1114 ACATCCCTGATTTGGAACCAAGAGGGGAGATTTCTATCCGTTATCCTTAAACAGGTACA 1173
Db 901 ACATCCCTGATTTGGAACCAAGAGGGGAGATTTCTATCCGTTATCCTTAAACAGGTACA 960
```

QY	1174	GCAATCCTAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAGGAGGATC	1233	Db	2041	TAATGGTTAGTATAGACAAACAGGGATCCTGTCCAGGGGCTTAAATTTGAAGC	2100
Db	961	GCAATNCCTAAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAGGAGGANN	1020	QY	2314	GCCTTAAATATAAATGGTCAATAGAAATGACCATGAGTGAAGACCTCCAATAT	2373
QY	1234	ATAGATGTCATAGATAAGGAACTAATCAACCTTTTGAAGATTCATTTGAAGAGTTGAA	1293	Db	2101	GCCTTAAATATAAATGGTCAATAGAAATGACCATGAGTGAAGACCTCCAATAT	2160
Db	1021	NN	1080	QY	2374	CTAGCTTCTCGATATGATTTTCACTTAGATGCTGTGGGCATCCAGGCTGGTCTAT	2433
QY	1294	TATATGCGAGCTGGATGACTCCTGAGGAAAAATATCTTGGTCCATCTACTAGAT	1353	Db	2161	CTAGCTTCTCGATATGATTTTCACTTAGATGCTGTGGGCATCCAGGCTGGTCTAT	2220
Db	1081	NN	1140	QY	2434	GGAGGATACCTCTCCCTGATGGCAATATGACAGAGTTCAGATATCTTCAGGGTTGCTATT	2493
QY	1354	CGCTCCGAGCTCGCTACAGATAGTTGATCTCACTGAAATTTATTTATCCAGTAGAA	1413	Db	2221	GGAGGATACCTCTCCCTGATGGCAATATGACAGAGTTCAGATATCTTCAGGGTTGCTATT	2280
Db	1141	CGCTCCGAGCTCGCTCCAGATAGTTGATCTCACTGAAATTTATTTATCCAGTAGAA	1200	QY	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTTTATGATACAGGATACACGGAACTGTATATG	2553
QY	1414	GATGATGTTATGGAAGCAGAGACTCATTTGAGTCACTGCTGATCTGTGAGCCACTA	1473	Db	2281	GCTGGGGCCCCAGTCACTCTGTGGATCTTTATGATACAGGATACACGGAACTGTATATG	2340
Db	1201	GATGATGTTATGGAAGCAGAGACTCATTTGAGTCACTGCTGATCTGTGAGCCACTA	1260	QY	2554	GGTCACCTCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2613
QY	1474	ATTATCTATGAAGAAACAAACAGATCTGGATAAATATCCATGATCTTTCATGTTTTT	1533	Db	2341	GGTCACCTCTGACAGAAATGAACANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2400
Db	1261	ATTANCAATGAAGAAACAAACAGATCTGGATAAATATCCATGATCTTTCATGTTTTT	1320	QY	2614	AAGTCCCTCTGGAACCAATCGTTTACTGCTTTACATGTTTCTCTGGATGAGAATGTC	2673
QY	1534	CCCAAGTCAAGAGAGGAAATTGAGTTTATTTTGGCTCTGATGTCGAAACAGGTTTC	1593	Db	2401	NN	2460
Db	1321	CCCAAGTCAAGAGAGGAAATTGAGTTTATTTTGGCTCTGATGTCGAAACAGGTTTC	1380	QY	2674	CATTTTGCACATACAGTATATTTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGAT	2733
QY	1594	CGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAACAAATATAAAGCATCGTGT	1653	Db	2461	CATTTTGCACATACAGTATATTTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGAT	2520
Db	1381	CGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAACAAATATAAAGCATCGTGT	1440	QY	2734	TTACAGATCTATCTCAGGAGACACAGCATACAGATTTCTGAACTGGGAGAACATTAT	2793
QY	1654	GGCTCGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTACCAGT	1713	Db	2521	TTACAGATCTATCTCAGGAGACACAGCATACAGATTTCTGAACTGGGAGAACATTAT	2580
Db	1441	GGCTCGCTGCTCCAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1500	QY	2794	GAATGATCTTTTGGACTACCTTCAGAAACCTTCGATCAGCTATGCTGCTCTAAAA	2853
QY	1714	GGTGAATGGGAAGTTCTTGCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG	1773	Db	2581	GAATGATCTTTTGGACTACCTTCAGAAACCTTCGATCAGCTATGCTGCTCTAAAA	2640
Db	1501	NN	1560	QY	2854	GTGATATAA 2862	
QY	1774	CTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833	Db	2641	GTGATATAA 2649	
Db	1561	CTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1620	RESULT 3			
QY	1834	TAGTAAATCTGGAGAGTGACAAGGCTGACTGACCTGCTGCTACTCACATTTCTTCTGC	1893	AK016546	3143 bp	mRNA	linear
Db	1621	TAGTAAATCTGGAGAGTGACAAGGCTGACTGACCTGCTGCTACTCACATTTCTTCTGC	1680	AK016546	Mus musculus adult male testis cDNA, RIKEN full-length enriched		HTC 03-APR-2004
QY	1894	ATCAGTCAGCATCTGACTCTTTTATAGTAAGTATAGTACCAAGAGAAATCCACTGT	1953	DEFINITION	Library, clone:4932434F09 product:DIPEPTIDYL PEPTIDASE 8 homolog		
Db	1681	ATCAGTCAGCATCTGACTCTTTTATAGTAAGTATAGTACCAAGAGAAATCCACTGT	1740	AK016546	[Homo sapiens], full insert sequence.		
QY	1954	GTGTCCCTTTACAGCTATCAAGTCTGAAAGTACCACCTGTCGAAACAAAGGAATTT	2013	ACCESSION	AK016546.1	GI:12855334	
Db	1741	GTGTCCCTTTACAGCTATCAAGTCTGAAAGTACCACCTGTCGAAACAAAGGAATTT	1900	VERSION	HTC; CAP trapper.		
QY	2014	TGGGCCACCAATTTTGGATTCAGCAGTCCCTTCTGACTATATCTCTCCAGAAATTTTC	2073	KEYWORDS	Mus musculus (house mouse)		
Db	1801	TGGGCCACCAATTTTGGATTCAGCAGTCCCTTCTGACTATATCTCTCCAGAAATTTTC	1860	SOURCE	Mus musculus		
QY	2074	TCCTTTGAAAGTACTCTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	2133	ORGANISM	Mus musculus		
Db	1861	TCCTTTGAAAGTACTCTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	1920	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
QY	2134	CAGCTCGAAAGAAATATCTACTGTGCTGTTTATATATGGTGGTCCCTCAGGTGAGTTG	2193	AUTHORS	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		
Db	1921	CAGCTCGAAAGAAATATCTACTGTGCTGTTTATATATGGTGGTCCCTCAGGTGAGTTG	1980	JOURNAL	Sciurognathi; Muridae; Murinae; Mus.		
QY	2194	GTGATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTCTCTAGGT	2253	PUBLISHED	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
Db	1981	GTGATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTCTCTAGGT	2040	REFERENCE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayaehizaki, Y.		
QY	2254	TATGTGTTGTAGTAGACACACAGGGATCTGTCCAGGGGCTTAAATTTTGAAGGC	2313	AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to		
					prepare full-length cDNA libraries for rapid discovery of new genes		
					Genome Res. 10 (10), 1617-1630 (2000)		
					11042159		
					3		
					Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
					Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,		

Qy	918	TGTGCAATAGCTAGCTAGCCAAATGGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTT	977	1998	CAAAACAAAGGAATTTTGGGCCACCATTTTCGATTTCAGAGGTCTCTTCTGACTATAC	2057
Db	1061	CGTACCAATAGCTAGCCAAATGGGAAGAGATGCCAGATCAGCTGGTGTGGCCACCTT	1120	2141	TAAACAAAGGAATTTTGGGCCACCATTTTCGATTTCAGAGGTCTCTTCTGACTACAC	2200
Qy	978	TGTTCTCCAAGAAGAAATTTGATAGATATCTGGCTATTGGTGGTGTCCAAAAGCTGAAAC	1037	2058	TCCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATGTTATGGGATGCTCTA	2117
Db	1121	TGTTCTTCCAAGAAGAAATTTGACAGATATCTGGCTACTGGTGGTGTCCCAAGCGGAAAG	1180	2201	CCCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACATGTTATGGGATGCTCTA	2260
Qy	1038	AACTCCCAAGTGGTGTAAAAATTTCTAGAAATCTATATGAAGAAAAATGATGAATCTGAGGT	1097	2118	CAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTCTGTTCATATATGTGG	2177
Db	1181	AACTCCTAGTGGTGTAAAAATTTCTAGAAATCTATATGAAGAAAAATGATGAATCTGAGGT	1240	2261	TAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGTTCATATATGTGG	2320
Qy	1098	GGAAATATTCATGTTACATCCCTATGTTGGAAACAAAGGAGGAGATTCATTCGGTTA	1157	2178	TCCTCAGGTGCTAGTTGGTGAATAATTCGGTTTTAAAGGAGTCAAGTATTTCCGCTTGAATAC	2237
Db	1241	GGAGATATTCATGTTACATCCCTATGTTGGAAACAAAGGAGGAGATTCATTCGGTTA	1300	2321	TCCCAGGTGCTAGTTGGTGAATAATTCGGTTTTAAAGGAGTCAAGTATTTCCGCTTGAATAC	2380
Qy	1158	TCCTAAACAGGTACAGCAATCTAAAGTCACTTTTAAGATGTCAGAAATATGATGA	1217	2238	CCTAGCCTCTCTAGGTTATGTTGTTAGTATGATAGACAAACAGGGGATCTCTGTCAACCGAGG	2297
Db	1301	TCCCAAAACAGGCACAGCAATCCCAAGGTCATTTTCAAGATGTCAGAAATATGATGA	1360	2381	CCTGGCCTCCCTGGGTTATGTTGTTAGTATGATAGACAAACAGGGGATCTCTGTCAACCGAGG	2440
Qy	1218	TGCTGAAGAGGATCATAGATGTCATAGATAGGAATTAATCAACCTTTTCAAGTCT	1277	2298	GCTTAAATTTGAAGGCGCTTTTAAATATATAAAATGGGTCAAAATAGAAATTCAGATCAAGT	2357
Db	1361	TGCTGCAGGAGGATTTAGATGTCATAGATAGGAATTAATCAACCTTTTCAAGTCT	1420	2441	ACTTAAATTTGAAGGCGCTTTTAAATATATAAAATGGGTCAAAATAGAAATTCAGATCAAGT	2500
Qy	1278	ATTGAAGAGGTTGAATATATTCGACAGCTGCCTAGAGTGGATCTCCTGAGGAAATATGCTTG	1337	2358	GGAGGACTCCCAATATCTAGCTTCGATATGATTTCACTTACCTTAGATCGTGTGGCAT	2417
Db	1421	GTTTGAGGAGTTGAATATATTCGACAGCTGCCTAGAGTGGATCTCCTGAGGAAATATGCTTG	1480	2501	GGAGGACTCCCAATATCTAGCTTCGATATGATTTCACTTACCTTAGATCGTGTGGCAT	2560
Qy	1338	GTCCATCTCTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAACCTGAAAT	1397	2418	CAAGCCTGCTCATGAGGATACCTCTCCCTGATGTCATTAATGAGGATTCAGAGGTCAGATAT	2477
Db	1481	GTCCATCTCTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAACCTGAAAT	1540	2561	CAAGCCTGCTCATGAGGATACCTCTCCCTGATGTCATTAATGAGGATTCAGAGGTCAGATAT	2620
Qy	1398	ATTATCCCAAGTGAAGATGATGTTATGGAAAGCAGAGACTCATTTAGTTCAGTGCCTGA	1457	2478	CTTACGGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATA	2537
Db	1541	ATTATCCCAAGTGAAGATGATGTTATGGAAAGCAGAGACTCATTTAGTTCAGTGCCTGA	1600	2621	CTTACGGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATA	2680
Qy	1458	TTCTGTGAGCGGCACCTAAATATCTATGAAGAAACAAAGACATCTGGATAAATATCCATGA	1517	2538	CACGGAACGTTATATGGGTCACTTCCCTGATGTCATTAATGAGGATTCAGTGTGGATCTGT	2597
Db	1601	CTCTGTGACCACTGATCATCTATGAAGAAACAAAGACATCTGGATAAATATCCACGA	1660	2681	CACGGAACGTTATATGGGTCACTTCCCTGATGTCATTAATGAGGATTCAGTGTGGATCTGT	2740
Qy	1518	CATCTTCAATGTTTTCCTCAAGTTCAGAGAGGAAATGAGTTTATTTTGGCTCTGA	1577	2598	GGCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGGTT	2657
Db	1661	TATTTTCAATGTTTTCCTCAAGTTCAGAGAGGAAATGAGTTTATTTTGGCTCTGA	1720	2741	GGCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGGTT	2800
Qy	1578	ATGCAAAACAGGTTTTCCTCAAGTTCAGAGAGGAAATGAGTTTATTTTGGCTCTGA	1637	2658	CCTGGATGAGAATGTCCATTTTGCACATACCAAGTATATTAAGTGTGTTTGTAGTGAGGC	2717
Db	1721	ATGCAAAACAGGTTTTCCTCAAGTTCAGAGAGGAAATGAGTTTATTTTGGCTCTGA	1780	2801	CTTGGATGAGAATGTCCATTTTGCACATACCAAGTATATTAAGTGTGTTTGTAGTGAGGC	2860
Qy	1638	TAAACGATCCAGTGGGCTGCTGCTCAGAGTGTTCAGTGTGCTTATCAAGAGGA	1697	2718	TGGAAGCCATATGATTTTACAGATCTATCCTCAGGAGAGACACAGCATAAGAGTTCTCTGA	2777
Db	1781	TAAACGATCCAGTGGGCTGCTGCTCAGAGTGTTCAGTGTGCTTATCAAGAGGA	1840	2861	TGGAAGCCATATGATTTTACAGATCTATCCTCAGGAGAGACACAGCATAAGAGTTCTCTGA	2920
Qy	1698	GATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAGT	1757	2778	ATCGGAGAACATTTAAAGTGAATTTTGAAGTCTGTAGAACTCTCTGTTATACACCTGG	2837
Db	1841	AATAACAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCTGGGT	1900	2921	GTCTGGAGAACATTTAAAGTGAATTTTGAAGTCTGTAGAACTCTCTGTTATACACCTGG	2980
Qy	1758	TGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCA	1817	2838	TATTGCTGCTTAAAGTGAATTTTGAAGTCTGTAGAACTCTCTGTTATACACCTGG	2897
Db	1901	TGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAAAGACTCTCTTTGGAGCATCA	1960	2981	TATCGTGTCTGAAAAGTGAATTAATCTCGACGCAATGTAGAACT---TGGTGTACACTG-	3036
Qy	1818	CCTGTACGTAGTCAAGTAAATCTGAGAGAGTGAAGGCTGACTGACCGTGGCTA	1877	2898	CTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGTATCATATTTT	2957
Db	1961	CCTGTACGTAGTCAAGTAAATCTGAGAGAGTGAAGGCTGACTGACCGTGGCTA	2020	3037	---TCTAACCAAAACAAAGAGGTTAGCCAAATGGAGAGCAGAGTGTATGAACCTACTTT	3093
Qy	1878	CTCACATCTCTGCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	1937	2958	GATACCTGCCATGATCAATCTACTCTGAAAAATAAATGTGGTGCATGC	3006
Db	2021	CTCACATCTCTGCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	2080	3094	TGTACCTGCCATGATCAATCTACTCTGAAAAATAAATGTGGTGCATGC	3142
Qy	1938	GGAGATCCACACTGTGTGCTCCCTTACAGCTATCAAGTCCCTGAGATGACCAACTTG	1997			
Db	2081	GAAGATCCACACTGTGTGCTCCCTTACAGCTATCAAGTCCCTGAGATGACCAACTTG	2140			

RESULT 4

AK029788

LOCUS

DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched

5517 bp

mRNA

linear

HTC 03-APR-2004

library, clone:4930560C15 product:DIPEPTIDYL PEPTIDASE 8 homolog
[Homo sapiens], full insert sequence.

ACCESSION AK029788
VERSION AK029788.1 GI:26081520
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
Riken integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 5517)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaiira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/

FEATURES
source
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers
1..5517
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:4930560C15"
/clone="taxon:10090"
/clone="4930560C15"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
2221..4837
misc_feature
/note="DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens]
(SPTR|Q9HBM5, evidence: FASTY, 95.7%ID, 100%length,
match=2649)
putative"

ORIGIN
Query Match 69.2%; Score 2158.6; DB 4; Length 5517;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 334; Indels 77; Gaps 6;
QY 201 AGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTGA 260
DB 2235 AGGAAGATGCAACATGCGCAGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTGA 2294
QY 261 AACTCGGAGCTGTGAGGAG---AATATTGAATCAGAGGATCGGCTTAATTTGGAGCCTTT 317
DB 2295 AACTGCGAGAGTGTGAGGAGGGAATGTTGAATCAGAGGACCGACCAACTGGAGCCTTT 2354
QY 318 TTATGTTGAGCGGTATTCTCTGGAGTCAGCTTAAAGAGCTGTTCGGCATACCAGAAAATA 377
DB 2355 TTATGTCGACGGTATTCTTGGAGTCAGCTGAAAGAGCTGCTGTGATACCAGAAAATA 2414
QY 378 TCATGGCTACATGATGGCTAAGGCACCATCATGATTTTCAATGTTTGTGAAGAGGAATGATCC 437
DB 2415 CCACGGCTACATGATGGCTAAGGCACCATCATGATTTTGTGAAGAGGACGATCC 2474
QY 438 AGATGGACCTCATTCAGACAGAAATCTATTACCTTGCCATGTCTGTGTGAGAACAGAGAAA 497
DB 2475 AGATGGCCTCACTCAGACAGGGTCTATTACCTTGCCATGTCTGTGTGAGAACAGAGAAA 2534
QY 498 TACACTGTTTATTCTGAAATTCCTAAAACCTATCAATAGACGAGCAGTCTTAATGCTCTC 557
DB 2535 TACACTATTTTATTCTGAAATTCCTAAAACCTATCAATAGACGAGCAGTCTTAATGCTCTC 2594
QY 558 TTGGAAGCCTCTTTTGGATCTTTTTCAGCAACACTGGCATATGGAATGATTTCTCGAGA 617
DB 2595 CTGGAAGCCTCTTTTGGATCTTTTTCAGCGCACCTAGACTATGGGATGATTTCTCGAGA 2654
QY 618 AGAAGAACTATTAAAGAGAAGAAACGCAATTCGGAACAGTCGGAATTTCTTACGATTA 677
DB 2655 GGAAGAGCTACTCAGAGAAGAAAGCGCAATTCGGAACGCTGGGATTCGCGATTTATGATTA 2714
QY 678 TCACCAAGGAAGTGGAACTATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAAGA 737
DB 2715 CCACCGGAAGCGGAACATTTCTGTTTCAAGCTGGCAGCGCATTTACCATATCAAGA 2774
QY 738 TGGAGGCCCAAGAGATTTACGCAACACTTTAAGGCCCAATCTAGTGGAACTAGTTG 797
DB 2775 TGGAGGACCAACATGGATTTTACACAGACTTTTGGGCCCAATTTAGTGGAACTAGTTG 2834
QY 798 TCCCAACATACGGAATGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTAT 857
DB 2835 TCCCAATATACGAATGATCCAAAATTTATGCCCTGCTGATCCAGACTGGATGCTTTTAT 2894
QY 858 ACATAGCAACGATTTTGGATCTTAACATCTTACCAGAGAGAAAGAGACTCACCTTA 917
DB 2895 TCACAGCAATGATTTTGGATATCAAACTCTGTAACAGGGAAGAGCGGAGATCACATA 2954
QY 918 TGTCCCAATGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTT 977

Tue Apr 18 08:18:38 2006

Db	2955	CGTACAAATGAGCTAGCCAACTGGAAGAGATCCAGATCAGCTGGTGGCCACTT	3014	3973	TAAACAAAGGAATTTTGGCCACCAATTTTGGATTGACAGGTCCTCTTCTGACTACAC	4032
Qy	978	TGTTCTCCAAAGAAATTTGATAGATTTCTGGCTATTTGGTGGTCCAAAGCTGAAAC	1037	2058	TCCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTACATGTTATGGGATCTCTCTA	2117
Db	3015	TGTTCTTCAAGAAATTTTCAAGATCTCTGGCTACTGGTGGTCTCCCAAGCGGAAG	3074	4033	CCCTCCAGAAATTTTCTTTTGAAGTACTACTGGTGGTCACTGATGGAATGTGTA	4092
Qy	1038	AACTCCCAAGTGGTGAATTTCTTGAATTTCTATATGAAGAAATGATGATCTGAGGT	1097	2118	CAAGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATGGTGG	2177
Db	3075	AACTCCTAGTGGTGAATTTCTTGAATTTCTATATGAAGAAATGATGATCTGAGGT	3134	4093	TAAAGCTCATGACCTACCACTGGAAAGAAATACCCCACTGTGTTATTCATATATGGTGG	4152
Qy	1098	GGAAATATTCATGTTACATCCCTATGTTGGAAACAAAGAGGGGAGATTCATTCGGTTA	1157	2178	TCCTCAGGTGAGTGGTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATAC	2237
Db	3135	GGAGATTAATTCATGTTACGTTCCCTATGTTGGAAACAAAGAGGGGAGATTCCTTTCGTTA	3194	4153	TCCCAGGTGAGTGGTGGTGAACAAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAACAC	4212
Qy	1158	TCCTAAACAGGTGACGCAATCTCTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATGA	1217	2238	CCTACCTCTCTAGGTTATGTGGTGTGATGATAGACAAACAGGGGATTCCTGTCCAGGAG	2297
Db	3195	TCCCAAAACA-----	3204	4213	CCTGGCTCTCTGGGTTATGTGGTGTGATAGACAAACAGGGGATTCCTGTCCAGGAG	4272
Qy	1218	TCCTGAAGGAGATCATAGATGATAGATAGGAATTAATTCAACTTTTGGATTTCT	1277	2298	GCCTAAATTTGAAGCGGCTTTTAAATATATAAATGGGTCAAATAGAAATTAAGCAAGGT	2357
Db	3205	-----GATTATAGATGATAGATTAAGAACTGGTTCAACCTTTTCGAGATTTCT	3252	4273	ACTTAAATTTGAAGCGGCTTTTAAATATATAAATGGGTCAAATAGAAATTCGATCAAGT	4332
Qy	1278	ATTGAGAGGTTGAATATATGTCAGAGCTGGATGACTCCTGAGGAAATATGCTTG	1337	2358	GGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGCAT	2417
Db	3253	GTTTGAGGAGTTGAATATATGTCAGAGCTGGATGACTCCTGAGGAAATATGCTTG	3312	4333	GGAGGACTCCAGTACCTAGCATCTCAGTATGACTTTCATTTGACTTGGATCGAGTGGGCAT	4392
Qy	1338	GTCCATCTCTAGATCGCTCCAGATCGGCTACAGATAGTGTGATCTCACTCGAAT	1397	2418	CMAGGCTGCTCTATGGGATACCTCTCCCTGATGGCATTAATGAGAGGTCAGATAT	2477
Db	3313	GTCCATCTCTAGATCGGTTCCAGATCTCAGATAGTGTGATCTCACTCGAAT	3372	4393	CMAGGCTGCTCTATGGGATACCTCTCCCTGATGGCATTAATGAGAGGTCAGATAT	4452
Qy	1398	ATTATTCAGGAGGATGATGTTATGGAAGGAGAGATCTATTTTGGTCTGCTG	1457	2478	CTTCAAGGTTGCTATTTGCTGGGCGGCGGATACCTCTCCCTGATGGCATTAATGAGAGGTCAGATAT	2537
Db	3373	ATTATTCAGGAGGATGATGTTATGGAAGGAGAGATCTATTTTGGTCTGCTG	3432	4453	CTTCCGGGTTGCTATTTGCTGGGCGGCGGATACCTCTCCCTGATGGCATTAATGAGAGGTCAGATAT	4512
Qy	1458	TTCTGTGAGGAGGATGATGTTATGGAAGGAGAGATCTATTTTGGTCTGCTG	1517	2538	CACGGAACGTTATATGGGTCACCTGACAGATGAGAGGCTATTTAGGATCTGT	2597
Db	3433	CTCTGTGAGGAGGATGATGTTATGGAAGGAGAGATCTATTTTGGTCTGCTG	3492	4513	CACGGAACGTTATATGGGTCACCTGACAGATGAGAGGCTATTTAGGATCTGT	4572
Qy	1518	CATCTTTTCATGTTTTCCTCAAACTCATGAAGTGAATTTGAGTTATTTTGGTCTG	3552	2598	GGCATGCAAGCAGAGAAAGTTTCCCTCTGAAACCAATCGTTTACTGCTTTTACATGTTT	2657
Db	3553	ATGCANAAAGGTTTTCCTCAAACTCATGAAGTGAATTTGAGTTATTTTGGTCTG	3612	4573	GGCATGCAAGCAGAGAAAGTTTCCCTCTGAAACCAATCGTTTACTGCTTTTACATGTTT	4632
Qy	1638	TAAACGATCCAGTGGTGGCTGCTCCTCAAGTGAATTTCAAGTGTCTCTATCAAAGAGGA	1697	2718	TGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGATTTCTGTA	2777
Db	3613	TAAACGATCCAGTGGTGGCTGCTCCTCAAGTGAATTTCAAGTGTCTCTATCAAAGAGGA	3672	4693	TGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGATTTCTGTA	4752
Qy	1698	GATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCGGATGGATCTAATATCCCAAGT	1757	2778	ATCGGAGAACATTTATGAACCTGATCTTTTGGCACTACCTTTCAAGAAACCTTTGGATCAG	2837
Db	3673	AATACCAATTTACAGTGGTGAATGGGAAGTTCTTGGCGGATGGATCTAATATCCCAAGT	3732	4753	GTCTGGAGAACATTTATGAACCTGATCTTTTGGCACTACCTTTCAAGAAACCTTTGGATCAG	4812
Qy	1758	TGATGAAGTCAAGAGGCTGGTATTTTGAAGGACCAAGATCTCCCTTTTAGAGCATCA	1817	2838	TATTTGCTCTTAAAGTGAATTAATTTTGAACCTGATTAATTTTGAACCTGATTAAGT	2897
Db	3733	TGATGAAGGACCAAGGCTGGTATTTTGAAGGACCAAGATCTCCCTTTTAGAGCATCA	3792	4813	TATTTGCTCTTAAAGTGAATTAATTTTGAACCTGATTAATTTTGAACCTGATTAAGT	4868
Qy	1818	CCTGTACGTAGTACGTAAATCTCTGAGAGGATGACAGGCTGATGACCTGGGCTA	1877	2898	CTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATTTGATCAACATTTT	2957
Db	3793	CCTGTACGTAGTACGTAAATCTCTGAGAGGATGACAGGCTGATGACCTGGGCTA	3852	4869	---TCTAAACCAACAGAGGTTTAAATCAACAGAAACACAGAAATTTGATCAACATTTT	4925
Qy	1878	CTCAATCTTTGCTGATCAGTCAAGTCTGATCTCTTTTAAAGTGAATTAAGTGAATTAAGT	1937	2958	GATACCTGCCATGTAACATCTATCTCTGAAATTAATTTGAGTGGTCCATGAGGGTCTACG	3017
Db	3853	CTCAATCTTTGCTGATCAGTCAAGTCTGATCTCTTTTAAAGTGAATTAAGTGAATTAAGT	3912	4926	TGTACCTGCCATGTAACATCTATCTCTGAAATTAATTTGAGTGGTCCATGAGAAATCTGCA	4985
Qy	1938	GAAGAAATCCACATGTGTGCTCTTTTAAAGTGAATTAAGTGAATTAAGTGAATTAAGT	1997	3018	GTTTTGTGG--TAGTAAATCTAATACCTTAAACCCCATGCTCAAAATCAATGATACATAT	3076
Db	3913	GAAGAAATCCACATGTGTGCTCTTTTAAAGTGAATTAAGTGAATTAAGTGAATTAAGT	3972	4986	GTTCATGATAGTACTCTAATACCTTAAAC---CATGCATTAATTAAGTGAATTAAGT	5041
Qy	1998	CAAAACAAAGGATTTTGGGCGCACATTTTGGATTCAGAGGTCCTCTTCTGACTATAC	2057	3077	CCTGAGAGACCCAGCAATACCAATAAGA	3103
				5042	CCCAAGGATTCAGACAGTACCTGAGGA	5068

```
RESULT 5
LOCUS AY411617
DEFINITION AY411617 2634 bp DNA linear GSS 16-DEC-2003
Mus musculus DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411617
VERSION AY411617.1 GI:39767585
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2634)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2634)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..2634
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2634
/gene="DPP8"
/locus_tag="HCM4283"
gene
ORIGIN
Query Match 63.6%; Score 1985; DB 10; Length 2634;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 446; Indels 21; Gaps 2;
QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTGAACTGGGACTGT 273
DB 1 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTGAACTGGGACTGT 60
QY 274 GAGGAG- - -AATATTGAATCAAGATCGGCTAAATTTGAGAGCTTTTATTTGAGCGG 330
DB 61 GAGGAGGGAATGGTGAATCAAGGACCGACCCAACTGGAGCTTTTATGTGGAACGG 120
QY 331 TATTCCTGAGTCAGCTTAAAGCTGCTCCCATACCAAGAAATATCATGCTGCTACATG 390
DB 121 TATTCCTGAGTCAGCTGAAAGCTGCTGCTGATACCAAGAAATACCAAGCTACATG 180
QY 391 ATGGCTTAGGCACCAATGATTTTCATGTTGTTGAAGAGGAATGATCCAGATGACCTCAT 450
DB 181 ATGGCTTAGGCACCAATGATTTTCATGTTGTTGAAGAGGAATGATCCAGATGACCTCAT 240
QY 451 TCAGACAGCAATCTATTACCTGTCATGCTGGTGAGAACAGAGAAATACATGTTTAT 510
DB 241 TCAGACAGGCTCTATTACCTGTCATGCTGGTGAGAACAGAGAAATACATGTTTAT 300
QY 511 TCTGAAATTCCTGAAATCTATCAATGAGCAGAGCTTATGCTCTCTTGGAGAGCTCTT 570
DB 301 TCTGAAATTCCTGAAATCTATCAATGAGCAGAGCTTATGCTCTCTTGGAGAGCTCTT 360
QY 571 TTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTA 630
```

```
Db 361 TTGGATCTTTTTCAGGCGACACTAGACTATGGATGTATTTCTCGAGAGAAAGAGCTACTC 420
QY 631 AGAGAAGAAACGCAATTCGAAACAGTCGGAATTTGCTTCTTACGATTTATCACAAGGAAGT 690
Db 421 AGAGAAGAAAGCGCAATTCGAAACCGTGGGGATCGCAGCTTTATGATTACCAAGGAGGAGC 480
QY 691 GGAACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAGTAAAGAGATGGAGGGCCACAA 750
Db 481 GGAACATTTCTGTTTCAAGCTGGCAGCGCATTTTACCATATCAAAAGATGGAGGACCAT 540
QY 751 GGAATTTACGCAACCAACCTTTTAAGGCCAATCTAGTGGAAACTAGTTGTGCCAACATACGG 810
Db 541 GGAATTTACACACAGCCTTTTGGGCCCAATTTAGTGGAAACTAGTTGTGCCAACATACGA 600
QY 811 ATGGATCCAAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGAT 870
Db 601 ATGGATCCAAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGAT 660
QY 871 ATTTGGATATCTAACTGCTAAACCGAGAGAAAGAGAGACTCCTATTTGTCGACATAGAG 930
Db 661 ATTTGGATATCTAACTGCTAAACCGAGAGAGAGAGAGATCACATACGTACACAATGAG 720
QY 931 CTAGCCAAACATGGAAGAAGATGCCAGATCCAGCTGAGTGCCTTCTTCTCCAGAA 990
Db 721 CTAGCCAAACATGGAAGAAGATGCCAGATCCAGCTGAGTGCCTTCTTCTCCAGAA 780
QY 991 GAATTTTGATAGATATTCTGGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050
Db 781 GAATTTTGACAGATATTCTGGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 1051 GGTAAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCAT 1110
Db 841 GGTAAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCAT 900
QY 1111 GTTACATCCCTATGTTGGAACAGAGGAGGAGATTTCAATCCGTTATCTTAAACACAGT 1170
Db 901 GTTACGTCCTCCCATGTTGGAACAGAGGAGGAGATTTCTTCTGTTATCTCCAAACAGGC 960
QY 1171 ACAGCAAACTCTAAAGTCACTTTTAAGATGTGAGAAATAATGATGATGCTGGAAGGAGG 1230
Db 961 ACAGCAAACTCTAAAGTCACTTTCAAGATGTGAGAAATAATGATGATGCTGGAAGGAGG 1020
QY 1231 ATCATAGATGTCATAGATAAGGAACTAAATTTCAACCTTTTGAGATTTCTATTTTGAAGAGGAT 1290
Db 1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080
QY 1291 GAATATATTGCCAGAGCTGGAGTCTCTGAGGGAATAATATGCTTGGTCCATCCTACTA 1350
Db 1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140
QY 1351 GATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTA 1410
Db 1141 GACCGTTCACAGACTCACCTGAGATAGTTCATGCTCCCTGAGTATTTATCCAGTA 1200
QY 1411 GAAGATGATGTTTGAAGAGGAGAGACTCAATGAGTCACTGCTGAGTCTGTCGAGGAGG 1470
Db 1201 GAAGATGATGCTATGAGAGAGAGAGACTTATAGAGTCACTGCTGAGTCTGTCGAGCA 1260
QY 1471 CTAATTTATCTATGAAGAAACAAACAGACATCTCGATTAATAATTCATGACATCTTTTCA 1530
Db 1261 CTGATCATCTATGAAGAAACAAACAGACATCTCGATTAATAATTCATGATTTTTCATGTT 1320
QY 1531 TTTCCCAAGAGTCAGAGAGGAAATGAGTTTATTTTGTCTGCTGAAATGCAAAACAGT 1590
Db 1321 TTTCTCAAACTCATGAAGATGAAATTTGAGTTTATTTTGTCTGCTGAAATGCAAAACAGT 1380
QY 1591 TTTCTGCTATTTATCAAAATTTATCATCTATTTTAAAGGAAAGCAAAATATAAAGAGTCCAGT 1650
Db 1381 TTTCTGCTATCTGATTAATAATTCATCCATTTTAAAGGAGAGCAAAATATAAAGAGTCCAGT 1440
QY 1651 GGTGGGCTGCTCTCCAAAGTATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTACC 1710
```


Db 249 CCGTAGTGGAAATTTATCACGTTAAAGATGGAGGCCACAGGATTTTACGCAACACCTT 308
Qy 770 TAAAGCCCAATCTAGTGGAAACTAGTGTCCCAACATACGAGTGAATCCAAAATTTATGCC 829
Db 309 TAAAGCCCAATCTAGTGGAAACTAGTGTCCCAACATACGAGTGAATCCAAAATTTATGCC 368
Qy 830 CCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCG 889
Db 369 CTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCG 428
Qy 890 TAACACAGAAAGAGAGACTCACTTATGTGCACAATGAGCTAGCCCAACATCGGAAGAAG 949
Db 429 TAACACAGAAAGAGAGACTCACTTATGTGCACAATGAGCTAGCCCAACATCGGAAGAAG 488
Qy 950 ATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAAAGAAATTTGATAGATATTCCTG 1009
Db 489 ATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAAAGAAATTTGATAGATATTCCTG 548
Qy 1010 GCTATTTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGTGTAATAATTTCTTAGAATTC 1069
Db 549 GCTATTTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGTGTAATAATTTCTTAGAATTC 608
Qy 1070 TATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTTACATCCCTATCTGTG 1129
Db 609 TATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTTACATCCCTATCTGTG 668
Qy 1130 AAACAAGGAGGCGAGATTCATTCGGTTATCTCTAAACAGGTACAGCAAAATCTTAAAGTCA 1189
Db 669 AAACAAGGAGGCGAGATTCATTCGGTTATCTCTAAACAGGTACAGCAAAATCTTAAAGTCA 728
Qy 1190 CTTTTAAGATGTCAGAAATATGATGTGCTGGAAGGAAGCATAGATGTCTATAGATA 1249
Db 729 CTTTTAAGATGTCAGAAATATGATGTGCTGGAAGGAAGCATAGATGTCTATAGATA 788
Qy 1250 AGGAATAATCAACCTTTTGAGATTCATTTTGAAGAGTGAATATTTGCCAGAGCTG 1309
Db 789 AGGAATAATCAACCTTTTGAGATTCATTTTGAAGAGTGAATATTTGCCAGAGCTG 848
Qy 1310 GATGAGCTCCTGAGGAAATATGCTGGTCCATCTACTAGATCGCTCCAGACTCGCC 1369
Db 849 GATGAGCTCCTGAGGAAATATGCTGGTCCATCTACTAGATCGCTCCAGACTCGCC 908
Qy 1370 TACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAA 1429
Db 909 TACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAA 968
Qy 1430 GGCAGAGACTCATGAGTCAGTGCCTGATCTCTGACGCCACTAAATATCTATGAAGAAA 1489
Db 969 GGCAGAGACTCATGAGTCAGTGCCTGATCTCTGACGCCACTAAATATCTATGAAGAAA 1028
Qy 1490 CAACAGACATCTGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTCAGGAAG 1549
Db 1029 CAACAGACATCTGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTCAGGAAG 1088
Qy 1550 AGGAAATTTAGTGTATTTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAA 1609
Db 1089 AGGAAATTTAGTGTATTTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAA 1148
Qy 1610 TTACATCTATTTTTAAGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAA 1669
Db 1149 TTACATCTATTTTTAAGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAA 1208
Qy 1670 GTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTC 1729
Db 1209 GTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTC 1268
Qy 1730 TTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAG 1789
Db 1269 TTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAG 1328
Qy 1790 GCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAAGTTCAGTTACGTAATTCCTGGAG 1849
Db 1329 GCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAAGTTCAGTTACGTAATTCCTGGAG 1388

Qy 1850 AGGTGACAAAGGCTGACTGACCGGTGGCTACTCAATCTTTGTGTCATCACTAGTCAGCACTGTG 1909
Db 1389 AGGTGACAAAGGCTGACTGACCGGTGGCTACTCAATCTTTGTGTCATCACTAGTCAGCACTGTG 1448
Qy 1910 ACTTCTTTTATAAGCTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGC 1969
Db 1449 ACTTCTTTTATAAGCTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGC 1508
Qy 1970 TATCAAGTCTGGAAGATGACCCCACTTGCAAAACAAAGAAATTTTGGGCCACCACTTTTGG 2029
Db 1509 TATCAAGTCTGGAAGATGACCCCACTTGCAAAACAAAGAAATTTTGGGCCACCACTTTTGG 1568
Qy 2030 ATTACAGAGTCTCTTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTA 2089
Db 1569 ATTACAGAGTCTCTTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTA 1628
Qy 2090 CTGGATTTACATTTGATGGATGCTCTACAAGCTCATGATCTACAGCTCGGAAGAAT 2149
Db 1629 CTGGATTTACATTTGATGGATGCTCTACAAGCTCATGATCTACAGCTCGGAAGAAT 1688
Qy 2150 ATCTACTGTGCTGTTTATATATGTTGTTCTCAGTGCAGTTGGTGAATAATCGTTTA 2209
Db 1689 ATCTACTGTGCTGTTTATATATGTTGTTCTCAGTGCAGTTGGTGAATAATCGTTTA 1722
Qy 2210 AAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGTTAGTGA 2269
Db 1723 ----- 1722
Qy 2270 TAGACAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAA 2329
Db 1723 ----- 1722
Qy 2330 TGGGTCAAAATAGAAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAG 2389
Db 1723 ----- 1722
Qy 2390 ATTTCAATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTCTATGAGGATACCTCTCCC 2449
Db 1723 ----- 1722
Qy 2450 TGATGGCATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCA 2509
Db 1723 -----GGTTGCTATTTGCTGGGGCCCCAGTCA 1748
Qy 2510 CTCTGTGGATCTCTATGATACAGGATACAGGAAGTTATATGGTCACTCCAGCCAGCA 2569
Db 1749 CTCTGTGGATCTCTATGATACAGGATACAGGAAGTTATATGGGTCACTCCAGCCAGCA 1808
Qy 2570 ATGAACAGGGCTATTTACTTTAGGATCTGTGGCCCATCGAAGCAGAAAAGTTCCCTCTGAAC 2629
Db 1809 ATGAACAGGGCTATTTACTTTAGGATCTGTGGCCCATCGAAGCAGAAAAGTTCCCTCTGAAC 1868
Qy 2630 CAAATCGTTTACTGTCTTTACATGTTTCTGGATGAGAATGTCCATTTTGCACATACCA 2689
Db 1869 CAAATCGTTTACTGTCTTTACATGTTTCTGGATGAGAATGTCCATTTTGCACATACCA 1928
Qy 2690 GTATATTTACTGAGTTTTTTAGTGGGCTGGAAAGCCATATGATTTTACAGATCTATCCTC 2749
Db 1929 GTATATTTACTGAGTTTTTTAGTGGGCTGGAAAGCCATATGATTTT -----C 1976
Qy 2750 AGGAGAGACACAGCATAAAGATTTCTGAATCGGGAAGACATATGAACATGCATCTTTTGC 2809
Db 1977 AGGAGAGACACAGCATAAAGATTTCTGAATCGGGAAGACATATGAACATGCATCTTTTGC 2036
Qy 2810 ACTACCTTCAAGAAAACCTTTGGATCACGTTATTTGCTCTCTAAAAGTATATAATTTTGC 2869
Db 2037 ACTACCTTCAAGAAAACCTTTGGATCACGTTATTTGCTCTCTAAAAGTATATAATTTTGC 2096
Qy 2870 CTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATAGGAGGTTTTTATCAACA 2929
Db 2097 CTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATAGGAGGTTTTTATCAACA 2156

QY	2930	GAACAACAGCAATTTGATCATCATCAATTTTGTATGCGATGCTTACAGCCTCATGATCTACA	2989	QY	2084	GTACTACTGATTTTACATTTGATTTGGATGCTCTACAGCCTCATGATCTACA	2140
Db	2157	GAACAACAGCAATTTGATCATCATCAATTTTGTATGCGATGCTTACAGCCTCATGATCTACA	2216	Db	241	GTACTACTGATTTTACATTTGATTTGGATGCTCTACAGCCTCATGATCTACA	300
QY	2990	TAAATGTGGTCCATGACAGGGCTACGGTTTGTGTAGTAACTTAATACCTTAACCCCA	3049	QY	2141	GAAGAATAATCTTACTGT- GCTGTTTATATATGTTGCTTCCAGGTGAGTTGGTGAAT	2199
Db	2217	TAAATGTGGTCCATGACAGGGCTACGGTTTGTGTAGTAACTTAATACCTTAACCCCA	2276	Db	301	GAAGAATAATCTTACTGTGNGGCTGTTTATATATGCGGGCCCTAAGTGAGTTGGTGAAT	360
QY	3050	CATGCTCAAAATCAAA 3065		QY	2200	AATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATAACCTAGCCTCTCTAGGTTATGTG	2259
Db	2277	CATGCTCAAAATCAAA 2292		Db	361	AATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATAACCTAGCCTCTCTAGGTTATGTG	420
RESULT 7				QY	2260	GTGTAGTAGTAGACAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCGCTTT	2319
AFI75225		1292 bp mRNA linear HTC 01-AUG-2003		Db	421	GTGNAGGGATAGACAACAGGGGATCCTGTCCCGAGGGCTTAAATTTGAAGCGCGCTTT	480
LOCUS		Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.		QY	2320	AAATATAAAATGGGTCAAAATAGAAATTAAGATCAGGTGGAAGGACTCCCAATATCTAGCT	2379
DEFINITION		AFI75225		Db	481	AAATATAAAATGGGTCAAAATAGAAATTAAGATCAGGTGGAAGGACTCCCAATATCTAGCT	540
ACCESSION		AFI75225.1 GI:33338055		QY	2380	TCTCGATATGATTTCAATTTGACTTTAGATCGTGTGGGCATCCACGCGCTGCTTATGGAGGA	2439
VERSION		HTC		Db	541	TCTCGATATGATTTCAATTTGACTTTAGATCGTGTGGGCATCCACGCGCTGCTTATGGAGGA	600
KEYWORDS		Homo sapiens (human)		QY	2440	TACCTCTCCCTGATGCGCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGTGGG	2499
SOURCE		Homo sapiens		Db	601	TACCTCTCCCTGATGCGCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGTGGG	660
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		QY	2500	GCCCCAGTCACTCTCTGATGCGCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGTGGG	2559
REFERENCE		1 (bases 1 to 1292)		Db	661	GCCCCAGTCACTCTCTGATGCGCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGTGGG	720
AUTHORS		Zhao, B., Xu, H. S., Tong, Y. K., Sheng, H., Qin, B. M., Liu, Y. Q., Liu, B., Wang, X. Y., Zhang, Q., Song, L., Gao, Y., Zhang, C. L., Ye, J., Ji, X. J., Liu, B. H., Lu, H., Chen, J. Z., Cai, M. Q., Zheng, W. Y., Teng, C. Y., Liu, Q., Yu, L. T., Lin, J., Gong, Q., Zhang, A. M., Gao, R. L. and Hui, R. T.		QY	2560	CTGACACAGATGAACAGGGCTTATTTAGTATGCTGTGGGCATCCAGGAGGAAAGTTC	2619
TITLE		Direct Submission		Db	721	CCTGACCAAGATGAACAGGGCTTATTTAGTATGCTGTGGGCATCCAGGAGGAAAGTTC	780
JOURNAL		Submitted (04-AUG-1999) Molecular Medicine Center for Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167, Bei Li Shi Lu, Beijing 100037, P.R. China		QY	2620	CCCTCTGAACAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTCATTTT	2679
FEATURES		source		Db	781	CCCTCTGAACAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGATGTCATTTT	840
		1. .1292		QY	2680	GCACATACCAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAGCCATATGTTTACAG	2739
		/organism="Homo sapiens"		Db	841	GCACATACCAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAGCCATATGTTTACAG	900
		/mol_type="mRNA" 9606"		QY	2740	ATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTTGAATCGGAGAAATTTGAACTG	2799
		/db_xref="taxon:9606"		Db	901	ATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTTGAATCGGAGAAATTTGAACTG	960
		/tissue_type="aorta"		QY	2800	CATCTTTTGACCTACCTTCAAGAAAAACCTTGGATCACGTTATTTGCTCTTAAAAAGTGATA	2859
		142. .1023		Db	961	CATCTTTTGACCTACCTTCAAGAAAAACCTTGGATCACGTTATTTGCTCTTAAAAAGTGATA	1020
		/product="MSTP135"		QY	2860	TAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGT	2919
		/protein_id="AAQ13650.1"		Db	1021	TAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGT	1080
		/tr_nrel="GI:33338056"		QY	2920	TAAATCAACAGAAACACAGAAATTTGATCATCATTTTGTATACCTGCGCATGTAACATCTTA	2979
		TSPHSTGLERNILXLYGPKVQVNNRFGVFFRNLTPFLGYVXGLDNKGS		Db	1081	TAAATCAACAGAAACACAGAAATTTGATCATCATTTTGTATACCTGCGCATGTAACATCTTA	1140
		CPRGLKFEAGFKPMQIIEIDDOVEGLQYLASRYDFDLDRVGHGWSYGLSLMAL		QY	2980	CTCTGAAAATAAATGTGGTGCCTACGAGGGTCTACGGTTTGTGTAGTAACTTAATAAC	3039
		MORSDFRVAIAGAPVTLWIFDYDTERYXMGHPDQNEQYILGSVAMQREKFPSEPN		Db	1141	CTCTGAAAATAAATGTGGTGCCTACGAGGGTCTACGGTTTGTGTAGTAACTTAATAAC	1200
		RLLLHGFDEVNHPAHTSILLFLVRAGKYDQLIYQERHSIRVPESGEHYELHLL		QY	3040	CTTAAACCCCATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCAT	3099
		HVLQENLGSRIAALKVI"		Db	1201	CTTAAACCCCATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCAT	1260
ORIGIN				QY	3100	AAGAATTTACTTAAAAAATAAAA 3120	
Query Match		39.0%; Score 1217.4; DB 4; Length 1292;		Db	1261	AAGAATTTCTTAAAAAATAAAA 1281	
Best Local Similarity		97.9%; Pred. No. 0;					
Matches 1254; Conservative		0; Mismatches 23; Indels 4; Gaps 2;					
QY	1844	CTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGATCATGTCAGC	1903	QY	1844	CTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGATCATGTCAGC	60
Db	1	CTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGATCATGTCAGC	60	QY	1904	ACTGTGACTCTTTTATAGTAACTAGTAAACAGAGAAATCCACACTGTGTCCCTTTT	1963
QY	1904	ACTGTGACTCTTTTATAGTAACTAGTAAACAGAGAAATCCACACTGTGTCCCTTTT	1963	Db	61	ACTGTGACTCTTTTATAGTAACTAGTAAACAGAGAAATCCACACTGTGTCCCTTTT	120
Db	61	ACTGTGACTCTTTTATAGTAACTAGTAAACAGAGAAATCCACACTGTGTCCCTTTT	120	QY	1964	ACAAGCTATCAGTCTTGAGATGACCCCACTTGGCAACAAAGGAAATTTTGGGCCACCA	2023
QY	1964	ACAAGCTATCAGTCTTGAGATGACCCCACTTGGCAACAAAGGAAATTTTGGGCCACCA	2023	Db	121	ACAAGCTATCAGTCTTGAGATGACCCCACTTGGCAACAAAGGAAATTTTGGGCCACCA	180
Db	121	ACAAGCTATCAGTCTTGAGATGACCCCACTTGGCAACAAAGGAAATTTTGGGCCACCA	180	QY	2034	TTTTGGATTACAGAGGTCCTCTCTGACTACTCTCCAGAAATTTCTCTTTTGA	2083
QY	2034	TTTTGGATTACAGAGGTCCTCTCTGACTACTCTCCAGAAATTTCTCTTTTGA	2083	Db	181	TTTTGGATTACAGAGGTCCTCTCTGACTACTCTCCAGAAATTTCTCTTTTGA	240

Li, H. B., Lu, H., Chen, J. Z., Cai, M. Q., Zheng, W. Y., Teng, C. Y.,
Li, Q., Yu, L. T., Lin, J., Gong, Q., Zhang, A. M., Gao, R. L. and Hu, R. T.
Homo sapiens normal aorta mRNA MST097

Unpublished
2 (bases 1 to 1278)
Xu, H. S., Zhao, B., Tong, Y. K., Sheng, H., Qin, B. M., Liu, Y. Q., Liu, B.,
Wang, X. Y., Zhang, Q., Song, L., Gao, Y., Zhang, C. L., Ye, J., Ji, X. J.,
Li, H. B., Lu, H., Chen, J. Z., Cai, M. Q., Zheng, W. Y., Teng, C. Y.,
Li, Q., Yu, L. T., Lin, J., Gong, Q., Zhang, A. M., Gao, R. L. and Hu, R. T.
Direct Submission
Submitted (30-JUL-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167
Bei Li Shi Lu, Beijing 100037, P.R. China

Location/Qualifiers
source
1..1278
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
1..1278
/gene="MST097"
591..1010
/gene="MST097"
/codon_start=1
/product="MSTP097"
/protein_id="AAQ13623.1"
/db_xref="GI:3338002"
/translation="MEIPSPDGIICRGQISSGAIAGPSHVSIMNOIHGRYMGHLTRNN
OGYLVGSVAMQAEKFPPEPNRLLLHGFLENVHFAHTSILLSFLVRAGKPYDLQIYP
QERHSIRVPSGEHVELHLLHQLNLSRIALKVI"

ORIGIN

Query Match 36.3%; Score 1132; DB 4; Length 1278;
Best Local Similarity 96.0%; Pred. No. 1.5e-302; Indels 11; Gaps 6;
Matches 1227; Conservative 0; Mismatches 40;
1844 CTGGAGGTGACAAAGGCTGACCGTGGCTACTACATTTCTTCTGTCATCAGTCAGC 1903
Db 1 CTGGAGGTGACAAAGGCTGACCGTGGCTACTACATTTCTTCTGTCATCAGTCAGC 60
Qy 1904 ACTGTGACTCTTTTATAAGTATAGTAAACGAAGATCCACACTGTGTCTCTTT 1963
Db 61 ACTGTGACTCTTTTATAAGTATAGTAAACGAAGATCCACACTGTGTCTCTTT 120
Qy 1964 ACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAACAAAGAAATTTGGGCCACCA 2023
Db 121 ACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAACAAAGAAATTTGGGCCACCA 180
Qy 2024 TTTTGGATTACAGAGTCTCTTCTCTGACTATCTCTCCAGAAATTTCTTTTGA 2083
Db 181 TTTTGGATTACAGAGTCTCTTCTCTGACTATCTCTCCAGAAATTTCTTTTGA 240
Qy 2084 GTACTACTGGATTACATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAA 2143
Db 241 GTACTACTGGATTACATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAA 300
Qy 2144 AGAATATCTACTGTCTGTTTCATATATGGTGGTCTCAGTGCAGTGTGTAATAATC 2203
Db 301 AGAATATCTACTGTCTGTTTCATATATGGTGGTCTCAGTGCAGTGTGTAATAATC 360
Qy 2204 GGTTTAAGAGTCAAGTATTCGGCTTGAATACCTAGCTCTCTAGGTATGTGGTTG 2263
Db 361 GGTTTAAGAGTCAAGTATTCGGCTTGAATACCTAGCTCTCTAGGTATGTGGTTG 420
Qy 2264 TAGTATAGACAAACAGGGATCTGTCCAGGCGCTTAANTTCAAGGCGCTTTAAAT 2323
Db 421 TAGTATAGACAAACAGGGATCTGTCCAGGCGCTTAANTTCAAGGCGCTTTAAAT 480
Qy 2324 ATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAAGGACTCCAAATATCTAGCTTCTC 2383
Db 481 ATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAAGGACTCCAAATATCTAGCTTCTC 540
Qy 2384 GATATG-ATTTCATTGACTTAGATCGTGGGCGCATCCACGGCTGGTCTATGGAGGATAC 2442

Db 541 GATATGAATTCATTGACTTAGATCGTGGGCAATCCCGCGTGTCTCTATGGAGTAC 600
Qy 2443 CTCTCCCTGATGGCAATTAATGCAGGTGAGATATCTTTCAGGGTTGCTATTTCTGGGCGC 2502
Db 601 CCTCTCCTGATGGCAT--ATGCAGAGTGCAGATATCTTTCAGG--TGCTATTTCTGGG--C 654
Qy 2503 CAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAACTTATATGGGTACACCT 2562
Db 655 CAGTCACTCTGTGATCTCTAT--GAACCCAGATACACGG--ACGTTATATGGGTACACCTG 710
Qy 2563 GACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCC 2622
Db 711 ACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCC 770
Qy 2623 TCTGAACCAAAATCGTTTACTCTCTTACATGTTTCTTGGATGAGATGTCCATTTTGCA 2682
Db 771 TTTGAACCAAAATCGTTTACTCTCTTACATGTTTCTTGGATGAGATGTCCATTTTGCA 830
Qy 2683 CATACAGTATATTACTGAGTTTCTTCTAGTGGGCTGGAAGCCCATATGATTACAGATC 2742
Db 831 CATACAGTATATTACTGAGTTTCTTCTAGTGGGCTGGAAGCCCATATGATTACAGATC 890
Qy 2743 TATCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTATGAATCTCAT 2802
Db 891 TATCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTATGAATCTCAT 950
Qy 2803 CTTTTCACATCCTTCAAGAAACCTTGGATCAGTATTTCTCTTAAAGTGATATAA 2862
Db 951 CTTTTCACATCCTTCAAGAAACCTTGGATCAGTATTTCTCTTAAAGTGATATAA 1010
Qy 2863 TTTTGACCTGTGTAGAACTCTCTGTATACATCTGGCTATTAAACAAATGAGGAGTTTA 2922
Db 1011 TTTTGACCTGTGTAGAACTCTCTGTATACATCTGGCTATTAAACAAATGAGGAGTTTA 1070
Qy 2923 ATCAACAGAAAAACAGAAATTCATCATCTTTTGTATACCTGGATGTAACATCTACTC 2982
Db 1071 ATCAACAGAAAAACAGAAATTCATCATCTTTTGTATACCTGGATGTAACATCTACTC 1130
Qy 2983 CTGAAATTAATGTGGTGCATCGAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTT 3042
Db 1131 CTGAAATTAATGTGGTGCATCGAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTT 1190
Qy 3043 AACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCAATACCATAAG 3102
Db 1191 AACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCAATACCATAAG 1250
Qy 3103 AATTACTTAAAAA 3120
Db 1251 AATTACTTAAAAA 1268

RESULT 10
BM557438
LOCUS
DEFINITION
AGENCOURT_6578992 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466759
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BM557438
BM557438.1 GI:18799430
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1042)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbe@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLC1967 row: k column: 16
 High quality sequence stop: 697.
 Location/Qualifiers
 1..1042
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5466759"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGCAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Query Match 30.1%; Score 938.2; DB 3; Length 1042;
 Best Local Similarity 98.7%; Pred. No. 1.1e-248;
 Matches 976; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

QY 1723 GAAGTCTTGGCGGCGATGGATCTAATATCCAAAGTTGATGAAGTCAGAGCGCTGGTATAT 1782
 DB 1 GAAGTCTTGGCGGCGATGGATCTAATATCCAAAGTTGATGAAGTCAGAGCGCTGGTATAT 60

QY 1783 TTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTACGTAAT 1842
 DB 61 TTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTACGTAAT 120

QY 1843 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGTGCATCAGTCAG 1902
 DB 121 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGTGCATCAGTCAG 180

QY 1903 CACTGTGACTCTTTTAAAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCCCTT 1962
 DB 181 CACTGTGACTCTTTTAAAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCCCTT 240

QY 1963 TACAAGTATCAAGTCTCAAGATGCCAAGTCCAAAGTTCGAAAGAAATTTGGGCCACC 2022
 DB 241 TACAAGTATCAAGTCTCAAGATGCCAAGTTCGAAAGAAATTTGGGCCACC 300

QY 2023 ATTTTGGATTACAGCGTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA 2082
 DB 301 ATTTTGGATTACAGCGTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA 360

QY 2083 AGTACTACTGGATTACATTGTATGGGATGCTCTCAAGCCCTCATGATCTACAGCCTGGA 2142
 DB 361 AGTACTACTGGATTACATTGTATGGGATGCTCTCAAGCCCTCATGATCTACAGCCTGGA 420

QY 2143 AAGAAATATCTACTGTCTGTTCATATATGTTGTCTCTCAGTGCAGTGGTGAATAAT 2202
 DB 421 AAGAAATATCTACTGTCTGTTCATATATGTTGTCTCTCAGTGCAGTGGTGAATAAT 480

QY 2203 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATCCCTAGCCTCTCTAGGTTATGTGTT 2262
 DB 481 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATCCCTAGCCTCTCTAGGTTATGTGTT 540

QY 2263 GTAGTGATAGACAAACAGGGGATCCCTGTACCAGGGGCTTAAATTTGAAGGCGCCTTTAAA 2322
 DB 541 GTAGTGATAGACAAACAGGGGATCCCTGTACCAGGGGCTTAAATTTGAAGGCGCCTTTANA 600

QY 2323 TATAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCT 2382
 DB 601 TATAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCT 660

QY 2383 CGATATGATTTCATTGACTTTAGATCTGTGGGATCCACGGCTGCTCTATGGAGTAC 2442
 DB 661 CGATATGATTTCATTGACTTTAGATCTGTGGGATCCACGGCTGCTCTATGGAGTAC 720

QY 2443 CTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTCCAGGTTTGTCTATTCTCTGGGCC 2502
 DB 721 CTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTCCAGGTTTGTCTATTCTCTGGGCC 780

QY 2503 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCT 2562
 DB 781 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCT 840

QY 2563 GACAGAATGAACAGGCTATTTACTTAGGATCTGTGGCCATCCAGACGAAAGTTCCTCC 2622
 DB 841 GACAGAATGAACAGGCTATTTACTTTANGATCTGTGGCCATCCAGACGAAAGTTCCTCC 900

QY 2623 TCTGAACCAAAATCGTTTACTGCTCTTACATG-GTTTCTGGATGAGATGT-CCATTTTG 2680
 DB 901 TCTGAACCAAAATCGTTTACTGCTCTTACATGCTTTCTGGATGAGATGTCCCATTTTG 960

QY 2681 CACATA-CCAGTATATTTACTGAGTTTTT 2708
 DB 961 CACATACCCAGGATATTTACTGGAGTTTTT 989

RESULT 11
 AK050021 3327 bp mRNA linear HTC 03-APR-2004
 LOCUS
 DEFINITION Mus musculus adult male liver tumor cDNA, RIKEN full-length
 enriched library, clone:C730003D12 product:DIPEPTIDYL PEPTIDASE 9
 homolog [Homo sapiens], full insert sequence.
 ACCESSION AK050021
 VERSION AK050021.1 GI:26340743
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 PUBMED 10349636

REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komio, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 PUBMED 11042159

REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Komio, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 PUBMED 11076861

REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation

JOURNAL REFERENCE AUTHORS	of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 3327) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayata, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishio, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasakawa, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, I., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissue was provided by William A. Held, Roswell Park Cancer Institute, Department of Molecular and Cellular Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers	Query Match 28.3%; Score 882.2; DB 4; Length 3327; Best Local Similarity 59.9%; Pred. No. 6e-233. Matches 1513; Conservative 0; Mismatches 1008; Indels 6; Gaps 2; 316 TTTTATGTTGACGGTATTCCTGGAGTCAGCTTAAAGCTGCTGGCGATACCAAGAA 375 Db 242 TTTCTGTGTCAGAGACATCTGTTGGATGGCTGCTGAGCTATATCCACGGCAGTCGAG 301 Qy 376 TATCATGGCTACATGATGGCTAAGGACCAATGATTTTCTGTTGTAAGAGGATGAT 435 Db 302 TCCTCGGGCTCATTTGTTCAGCAAGGCCCCACACACTTCCAGTTTGTGCAAGAGCTTCGAC 361 Qy 436 CCAGATGGACCTCATTTTCAGACAGAGATCTTATACCTTGGCCATGTCTGTGTGAGAAACAGAGAA 495 Db 362 GAGTCTGGCCCCCACTCTCACCGTCTCTATTTACCTTCGGAATGCCCTTACGCGACGCCGTGAG 421 Qy 496 ATATACACTGTTTATTTCTGAAATTCCTCAAAATCTCAATAGACAGCAGCTCTTAATGCTCTC 555 Db 422 AACTCTCTCTCTACTTCGAGATCCCAAGAAAGTGGCAAGAGAGGCCCTGCTGCTGCTG 481 Qy 556 TCTTGAAGCTCTTTTGGATCTTTTTCAGGCACACTGGAGCTATGGAATGATTTCTCGA 615 Db 482 TCCTGGAAGCAGATGCTGACCACTTCCAGGCCACACCCACCATGGTGTCTACTCCCGA 541 Qy 616 GAAGAAGACTATTAAAGAGAAAGAAAGCGCATTCGACAGTCGGAATTTGTTCTTACGAT 675 Db 542 GAGAGAGCTACTCGGGAGCGCAAGCGCTGGGCGTCTTCGGAATCACCTCTTATGAC 601 Qy 676 TATCACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGAATTTATCAGGTAA 735 Db 602 TTCACAGTGAAGCGGCTCTTCTTTCAGGCCAGCATAGCTGTTCCTCCTGACG 661 Qy 736 GATGGAGGCGCACAGGATTTACGCAACAACTTTTAAAGGCCAATCTAGTGAAGTACT 795 Db 662 GATGGTGGCAAGAAATGGCTTTATGTGTCCCGATGAAGCCCATCGAGATCAAGACTCAG 721 Qy 796 TGTCCCAACATACGAGTGGATCCAAATTTATGCCCGGTGATTCAGACTGGAATGCTTTT 855 Db 722 TGTCTGGGCGACGATGGACCCCAAAATCTGCCCGGAGACCTTCTCTTTTCTTC 781 Qy 856 ATACATAGCAACGATATTTGGATATCTAAATCTGTAACAGAGAGAAAGAGACTACT 915 Db 782 ATCAACAAACAGTATCTGTGGGTGGCAACATCGAGACTGGGGAGGAACGGCGGTCA 841 Qy 916 TATGTGCACAAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACC 975 Db 842 TTCTGTCAACAGGTTTCAGCTGTGTCTCTGGACAATCCCAATCAGCAGGGGTGGCCACC 901 Qy 976 TTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGTGTCCTCAAAAG-CT 1032 Db 902 TTTGTCTATCAGAGAGGATTCGACCGCTTCACTGGGTGCTGTGTGTCGCCACGGCTCT 961 Qy 1033 GAAACAACCTCCAGTGGTGAATTTCTAGAAATCTTATATGAAGAAATGATGAATCT 1092 Db 962 TGGGAAGGCTCCGAAGGTCTCAAGACGCTGCGCATCTTATATGAGAAAGTGGACGATCT 1021 Qy 1093 GAGTGGAAATTTATTTATGTTATATCTTATATCTTATGTTGGAACAGAGGCGGATTCATTC 1152 Db 1022 GAGTGGAGGTTCATTTATGTTGCTTCCCGCTCCCGCTGGAGAGAGAGAGGACTCTTAC 1081 Qy 1153 CGTTATCTTAAACAGGTAACAGAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATG 1212 Db 1082 CGCTACCCAGGACAGCAGCAAGAACCCCAAGATTCGCTGAGCTGCTGCTGCTCCAG 1141 Qy 1213 ATTGATGCTGAAGAGGATCATAGATGTATAGATGATAGAGAACTAACTTCAACCTTTGAG 1272 Db 1142 ACGGACCATCAGGCCAAATTCGTGTCGAGTGGAGAGAACTGGTACAGGCATTCAGC 1201 Qy 1273 ATTCTATTGAAGAGTGAATATATTTGCGCAGAGCTGGATGGAATCTCTGAGGGAAATAT 1332 Db 1202 TCCCTTTTCCCAAAAGTGGAGTATCGCCCGGCTGGCTGGACACGGGACGGCAATAT 1261
TITLE JOURNAL	of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 3327) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayata, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishio, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasakawa, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, I., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissue was provided by William A. Held, Roswell Park Cancer Institute, Department of Molecular and Cellular Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers	Query Match 28.3%; Score 882.2; DB 4; Length 3327; Best Local Similarity 59.9%; Pred. No. 6e-233. Matches 1513; Conservative 0; Mismatches 1008; Indels 6; Gaps 2; 316 TTTTATGTTGACGGTATTCCTGGAGTCAGCTTAAAGCTGCTGGCGATACCAAGAA 375 Db 242 TTTCTGTGTCAGAGACATCTGTTGGATGGCTGCTGAGCTATATCCACGGCAGTCGAG 301 Qy 376 TATCATGGCTACATGATGGCTAAGGACCAATGATTTTCTGTTGTAAGAGGATGAT 435 Db 302 TCCTCGGGCTCATTTGTTCAGCAAGGCCCCACACACTTCCAGTTTGTGCAAGAGCTTCGAC 361 Qy 436 CCAGATGGACCTCATTTTCAGACAGAGATCTTATACCTTGGCCATGTCTGTGTGAGAAACAGAGAA 495 Db 362 GAGTCTGGCCCCCACTCTCACCGTCTCTATTTACCTTCGGAATGCCCTTACGCGACGCCGTGAG 421 Qy 496 ATATACACTGTTTATTTCTGAAATTCCTCAAAATCTCAATAGACAGCAGCTCTTAATGCTCTC 555 Db 422 AACTCTCTCTCTACTTCGAGATCCCAAGAAAGTGGCAAGAGAGGCCCTGCTGCTGCTG 481 Qy 556 TCTTGAAGCTCTTTTGGATCTTTTTCAGGCACACTGGAGCTATGGAATGATTTCTCGA 615 Db 482 TCCTGGAAGCAGATGCTGACCACTTCCAGGCCACACCCACCATGGTGTCTACTCCCGA 541 Qy 616 GAAGAAGACTATTAAAGAGAAAGAAAGCGCATTCGACAGTCGGAATTTGTTCTTACGAT 675 Db 542 GAGAGAGCTACTCGGGAGCGCAAGCGCTGGGCGTCTTCGGAATCACCTCTTATGAC 601 Qy 676 TATCACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGAATTTATCAGGTAA 735 Db 602 TTCACAGTGAAGCGGCTCTTCTTTCAGGCCAGCATAGCTGTTCCTCCTGACG 661 Qy 736 GATGGAGGCGCACAGGATTTACGCAACAACTTTTAAAGGCCAATCTAGTGAAGTACT 795 Db 662 GATGGTGGCAAGAAATGGCTTTATGTGTCCCGATGAAGCCCATCGAGATCAAGACTCAG 721 Qy 796 TGTCCCAACATACGAGTGGATCCAAATTTATGCCCGGTGATTCAGACTGGAATGCTTTT 855 Db 722 TGTCTGGGCGACGATGGACCCCAAAATCTGCCCGGAGACCTTCTCTTTTCTTC 781 Qy 856 ATACATAGCAACGATATTTGGATATCTAAATCTGTAACAGAGAGAAAGAGACTACT 915 Db 782 ATCAACAAACAGTATCTGTGGGTGGCAACATCGAGACTGGGGAGGAACGGCGGTCA 841 Qy 916 TATGTGCACAAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACC 975 Db 842 TTCTGTCAACAGGTTTCAGCTGTGTCTCTGGACAATCCCAATCAGCAGGGGTGGCCACC 901 Qy 976 TTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGTGTCCTCAAAAG-CT 1032 Db 902 TTTGTCTATCAGAGAGGATTCGACCGCTTCACTGGGTGCTGTGTGTCGCCACGGCTCT 961 Qy 1033 GAAACAACCTCCAGTGGTGAATTTCTAGAAATCTTATATGAAGAAATGATGAATCT 1092 Db 962 TGGGAAGGCTCCGAAGGTCTCAAGACGCTGCGCATCTTATATGAGAAAGTGGACGATCT 1021 Qy 1093 GAGTGGAAATTTATTTATGTTATATCTTATATCTTATGTTGGAACAGAGGCGGATTCATTC 1152 Db 1022 GAGTGGAGGTTCATTTATGTTGCTTCCCGCTCCCGCTGGAGAGAGAGAGGACTCTTAC 1081 Qy 1153 CGTTATCTTAAACAGGTAACAGAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATG 1212 Db 1082 CGCTACCCAGGACAGCAGCAAGAACCCCAAGATTCGCTGAGCTGCTGCTGCTCCAG 1141 Qy 1213 ATTGATGCTGAAGAGGATCATAGATGTATAGATGATAGAGAACTAACTTCAACCTTTGAG 1272 Db 1142 ACGGACCATCAGGCCAAATTCGTGTCGAGTGGAGAGAACTGGTACAGGCATTCAGC 1201 Qy 1273 ATTCTATTGAAGAGTGAATATATTTGCGCAGAGCTGGATGGAATCTCTGAGGGAAATAT 1332 Db 1202 TCCCTTTTCCCAAAAGTGGAGTATCGCCCGGCTGGCTGGACACGGGACGGCAATAT 1261
FEATURES	source	1..3327 /organism="Mus musculus" /mol_type="mRNA" /strain="CE7BL/6J" /db_xref="PANTOM:DB:C730003D12" /db_xref="taxon:10090" /clone="C730003D12" /sex="male" /cell_type="tumor" /tissue_type="liver" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 176..2764 /note="unnamed protein product; DIPEPTIDYL PEPTIDASE 9 homolog [Homo sapiens] (SPTR AAL47179, evidence: FASTY, 92%ID, 99.6%length, match=2580) putative" /codon_start=1 /protein_id="BAC34034.1" /db_xref="GI:26340744" /translation="MCSGVSPVEQVAAGMDMDTAAFCVQKSHWDGLRSIHGSRKSS GLIVKAPHDFOVQKPDGSPHSHRLYLGMYPGSRNSLLYSEIPKPKVRKEALLLL SWKQMDHPQATPHGVYSREBELLRKRLGVFTSYDFHSGSLFLFOANSFLH CRDGGKNGFMVPMKPLEIKTCQSGRMDPKICPADPAFESFINNSDLVANIETGEE RLTFCQGSAGVLDNPKSAGVATFVIEEDRFTGCWCPCTASWEGSEGLKRLILY KELVQSFSLPKVSVIARAGTRDGVAMFLDRPQRLQLVLLPPLFIVSESE ACHQRAARVQKVPFVIEVNTVINVDIHPFPQAEQQDQCFCLANECKTGF CHLYRVTEKTKVDWTEPLSTDEPKFRLTEVALTSGEMVLSRHGSKIWNVEQ TKLAVYQGTQKTDPLEHLVSVESAGIRVLTLLTGFHSCMSQSPDMFVSHVSVS TPCVHYKLSPEDDDPLHKQPFVSWMEANCPDVPVPEIPFHTRDVLYGMI YKPHITLQPKRQPTLVLYGQVQVQVNNPKGIKYLPLNTLASLGVAVVIVDGRSC QRGHFEAGALQMGQVEIEDQGLQVASKYGFIDLSRVAHGWYGGFSLMGLI HKQFVKVIAIGAPVYVMAYDTGTEARMVDPNNQGVAGSVLHVKLPENR LLIHGLFDLENVHFHTNPLNQLIRAGKPVQLQIYPNERHSIRCSGEHYEVLHH FLOEHL"
CDS		
ORIGIN		

QY	1333	GCTGGTCCATCTACTAGATCGTCCAGACTCGCTACAGATAGTGTGATCTCACCT	1392	Db	2342	GTCCCATCCATGGCTCGTCTACGGGGCTTCTCTCTCATCTATGGGGCTCATCCACAG	2401
Db	1362	GCCTGGGCAATGTTCTTGGGACCGTCCCGCAACCGCTTGTCTCTCTCGCCCT	1321	QY	2470	TCAGATATCTTCAAGGTTGCTATTGTCTGGGGCCCGCAGTCACTCTGTGGATCTTCTATGAT	2529
QY	1393	GAAATATTTATCCAGTAGAAGATGATGTTATGAAAGGCAGAGACTCATTTGAGTCAAGT	1452	Db	2402	CCACAAGTGTTCRAAGGTAGCCATTGCGGGCGCTCTGTCTCACTGTGTGGATGGCTTATGAC	2461
Db	1322	GCTCTCTTCATCCCGCGGTGAGAGTAGGGCCCGCAGCGAGCTGCCAGAGCCGTC	1381	QY	2530	ACAGGATACACGGAACTTATATGAGGTCACTCCCTGACAGAGTGAACAGGGCTTATCTT	2589
QY	1453	CTGTGATCTGTGACGCCCAATTTATCTATGAAGAAACACAGACATCTGGATAAATATC	1512	Db	2462	ACAGGTACACGGAAAGATACATGGATGTCCTCCGAAATTAACAGCAAGGCTATGAGCA	2521
Db	1382	CCCAAGATGTGCAGCCCTTGTCTATGAAAGAGTCAACCAATGTCTGGATCAAGTC	1441	QY	2590	GGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATGTTTACTGCTCTT	2649
QY	1513	CATGACATCTTTCTATGTTTTCCTCCAAAGTCA---CGAAGAGAAATTTAGTTTATTTT	1569	Db	2522	GGGTCTGTAGCCCTGTGATGTGAGAGAGTGCCTCAATGAGCTTAACCGCTGTCTATCTC	2581
Db	1442	CACGACATCTTCCACCCGTTTCTCTCAGGCTGAGGGCCAGCAGACTTTTGTCTCTCGT	1501	QY	2650	CATGTTTCTCTGATGAGAAATGTCATTTTTCGACATACCAGTATATTTATGAGTTT	2709
QY	1570	GCTCTGAATCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGAA	1629	Db	2582	CACGGCTTCTGGACGAGACGTTCTACTTCTCCACACAAATTTCTGTGTGCCAGCTG	2641
Db	1502	GCCAAAGAAATGCAAGACTGGCTTCTGACCTGTACAGGTCAAGTCAAGTGAACCTTAAC	1561	QY	2710	GTGAGGGCTGAAAGCCATATGATTTACAGATCTATCTTCAGAGAGACACAGCAAGA	2769
QY	1630	AGCAATATATAACGATCCAGTGTGGCTGCTCTCAAGTGATTTCAAGTGTCTTATC	1689	Db	2642	ATCCGAGCAGGAAAGCCATACACAGCTTCAGATCTACCCAAACGAGAGACATAGCATCCG	2701
Db	1562	AGGACTATGACTGACGGAACCCCTCAGCCCTACAGAGATGAGTTTAAGTGCCTCATC	1621	QY	2770	GTTCTGATCGGAGAACATTTATGATGATCTTTTTCGACATCTTTTCGACATCTTCAAGAAACCT	2829
QY	1690	AAAGAGGATAGCAATTTACAGTGTGAAATGGAAAGTTCTTGGCCGGGATGATCTAT	1749	Db	2702	TGCGCGAGTCCGAGAGCATTTACGAGTGAAGTGTGCTGCTCTTCTGAGGAACACCTG	2761
Db	1622	AAGGAGGAGTGCCTCTGACAGTGGCGAGTGGAGGTCTTTCGAGGCGATGGCTCCAAG	1681	QY	2830	GGATCAC 2836	
QY	1750	ATCCAGTGTGATGAGTCAAGAGCTGTGATATTTTGAAGGCACCAAGAGCTCCCTTTA	1809	Db	2762	TGACCTC 2768	
Db	1682	ATCTGGGTCAACGAGCAGAGCAAGCTGGTGTACTTTCAAGGTACAAAGCACACCGCTG	1741	RESULT 12			
QY	1810	GAGCATCACTGTACGTAGTCAAGTTCAGTAAATCTCGAGAGGTGACAAAGCTGACTGC	1869	AK078301			
Db	1742	GAAATCACTCTATGTGTGCTCAGTACAGTCAAGGCGAGATCGTGGCTCACACG	1801	LOCUS			
QY	1870	GTGGTACTCACATCTTGTGATCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	1929	DEFINITION			
Db	1802	CTGGCTTCTCCACAGCTGCTCCATGAGCCAGAGCTTCGACATGCTGAGTCACTAC	1861	ACCESSION			
QY	1930	AGTACACAGAGATCCACACTGTGTGCTCTTTTACAGCTATCAAGTCTCAAGTCAAGT	1989	VERSION			
Db	1862	AGCAGTGTGACACGCAACCTCTGTGTATGTATGATGATCAAGTCAAGGCGCCGATGAT	1921	KEYWORDS			
QY	1990	CAACTTGCACAAAGGAATTTTGGGCCACCATTTTGGATTTCAAGAGTCTCTCTCT	2049	SOURCE			
Db	1922	CCATGACACAGCAACACGCTTCTGGCCAGCATGATGAGGAGCCCAATGCCCCCA	1981	ORGANISM			
QY	2050	GACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATGATGATGG	2109	REFERENCE			
Db	1982	GACTATGTCCTGAGATCTTCCACTTCCACACCCGTGACAGCTGACGCTCTACGCG	2041	AUTHORS			
QY	2110	ATGCTCTACAGCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTCTATA	2169	TITLE			
Db	2042	ATGATCTCAAGGCACACACCTGCAACCTGGAGGAGCAACCCCACTGTCTCTTGTG	2101	JOURNAL			
QY	2170	TATGTGTCTCTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	2229	PUBMED			
Db	2102	TATGGGGCCCAAGGTGCGATGTTGTTGATCACTCTTTTAAAGGATCAATATCTGCGG	2161	REFERENCE			
QY	2230	TTGAATACCTCTAGCTCTCTAGGTATGATGATGATGATGATGATGATGATGATGAT	2289	AUTHORS			
Db	2162	CTAATACACTGGCATCTTGGGCTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2221	TITLE			
QY	2290	CACGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATAATGGTCAATAGAAATGAC	2349	JOURNAL			
Db	2222	CAGCGGGCTCTGCTTCAAGGGGGCCCTGAAATCAATGGGCGAGTGGAGATTGAG	2281	PUBMED			
QY	2350	GATCAGGTGGAGGACTCCATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGT	2409	REFERENCE			
Db	2282	GACCAAGGTGGAGGCTTGCAGTACGTGGCTGAGAGATGAGCTTCAATGACTTGAAGCG	2341	AUTHORS			
QY	2410	GTGGGCATCCACGGCTGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGAGAGG	2469				

AK078301 3457 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male olfactory brain cDNA RIKEN full-length
enriched library, clone:6430584G11 product:DIPEPTIDYL PEPTIDASE 9
homolog [Homo sapiens], full insert sequence.

AK078301 GI:26347124
HTC; CAP trapper.
Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.

Tue Apr 18 08:18:38 2006

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3457)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, I., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

1. 3457

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:6430584G11"

/db_xref="taxon:10090"

/clone="6430584G11"

/sex="male"

/tissue type="olfactory brain"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

182_-2770

/note="unnamed protein product; DIPEPTIDYL PEPTIDASE 9 homolog [Homo sapiens] (SPTRLAAL47179, evidence: PASTY, putative)"

/codon_start=1

/protein_id="BAC37211.1"

/db_xref="GI:26347125"

/translation="MGSGVSPVEQVAGDMDDTAARFCVQKHSWDGLRSIIHGRKSS GLVSKAPHDFQVQPDGSGHSHLYLGMYPGSRNSLLYSEIPKKVRKEALLLL SWKMLDHFQTHGVGVSREELRLERKRLGVFGITSYDFHSESLPLFOANSFLH CRDGGKFWSPKPLEIKTQCGRPMDPKICPADPAPFSPINNSDLVWANIETGEE RRLTFCHQSGAGVLDNPKSGVATVUIQEPDFTGWCWCPSTASWESSEGLKTLRLY BEVDESEVTHVSPALEREKIDSVYBRTGSKPKIALKLAELQTDHQKIVSSCE KELVQFSSLLPFVKEYIARAGNRTRDKYANMFLYRPFQASQDDPCFLRANECKTGF AQRAARAVPKNVQPEVIYEVINVINVINHDIHPHFPQASQDDPCFLRANECKTGF CHLRYVTEKTDYDTEPLESPTEDEFKCIKEVALTSGEMVELSRHGSKIVVNEQ TKLVYFOQKTDPLEHLYLVVSVESAGEIVRLITLGFSSCSMSQSPMPFVSHYVNS TPCCVHYVTLGSGDDPLHKQPRFASMANCNPDIYYPPEIHPHTFRAVDQLYGM IYKPHTLQPRKHQNTQVIEIDVQGVQVLNNSFKIGLIRLNLILASLCYAVVLDGRSC QRGLHLPQKXNQMGQVEIEDVQGVQVLAERQYFKIDLSRVALHWSYGGSLMGLI HKPQVFKVAGAVFTVMAYDVTGYTERYMDVPENNKQYEGAGSVLHVEKLHPNPR LLIHLGFLDENVHFHTNFLVSQLRAGKPYQLQIYPNERHSIRCRSEGEHYEVTLLH FLQEHL"

ORIGIN

Query Match 28.2%; Score 880.6; DB 4; Length 3457;
Best Local Similarity 59.8%; Pred. No. 1.7e-232;
Matches 1512; Conservative 0; Mismatches 1009; Indels 6; Gaps 2;

QY 316 TTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAGAGCTGCTTGGCGATACACAGAAAA 375
DB 248 TTCTGTGTGCGAAGACACTCGTGGGATGGCTGCTAGCAATATCCACGGCAGTGCAG 307

QY 376 TATCATGGCTACATGATGGCTAAAGGCACCAATGATTTTCATGTTGTGAGGAAGATCAT 435
DB 308 TCCTCGGCTCATTTGTGACCAAGGCCCCAGCACTTCCAGTTGTGTGCAAGAGCTGAC 367

QY 436 CCAGATGAGCACTCAATTCAGACAGAAATCTATTACCTTGGCCATGTCTGGTGAGAACAGAA 495
DB 368 GAGTCTGGCCCCCACTCTCCACCGTCTCTATTACCTCGGAATGCTTTACCGCAGCGTGAG 427

QY 496 AATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGACAGACTCTTAATGCTC 555
DB 428 NAATCCCTCTCTACTCCGAGATCCCAAGAAAGTGGGAAGGAGGCCCTGCTGCTGCTG 487

QY 556 TCTTGAAGCCCTCTTTTGGATCTTTTTCAGGCACCACTGGACTATGGAATGTATTCTCGA 615
DB 488 TCTTGAAGCAGATGCTGGACCACTTCAGGCCACACCCACCACCTGCTACTCCCCGA 547

QY 616 GAAGAAGACTATTAAAGAGAAGAAAAAGCAATTTGGAACAGTTCGGAATTTGCTTCTTACGAT 675
DB 548 GAGGAGGAGCTACTCGCGGAGCGCAAGCGCTGGGCGTCTTCGGAATCACTCTTATGAC 607

QY 676 TATCACCAGGAAGTGGAAACATTTCTTCAAGCCGTAGTGGAAATTTATACGTAATAA 735
DB 608 TTCACAGTGAAGCGGCGCTCTTCTCTTCAGGCCAGCAATAGCTGTGTTCCTACTGCAGG 667

QY 736 GATGAGGCGCCAAAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGT 795
DB 668 GATGGTGGCAAGATGGCTTTATGTGTCCCGATGAAGCACTGGAGATCAAGACTCAG 727

QY 796 TGTCCTCAACATACGAGTGGATCCAAATATGTCCTCCCTGATCAGACTGGATTTGCTTTT 855
DB 728 TGTCTGGGCGACGATGAGACCCCAAAATCTGCCCCGACACCTGCTCTTTTCTCTTC 787

QY 856 ATACATAGCAACGATATTTGGATATCTTAACATGTGTACAGAGAAAGAGAGACTCACT 915
DB 788 ATCAACAACAGTATCTGTGGTGGCAACATCTGAGACTGGGAGGAACGGCGCTCACC 847

QY 916 TATGTGCACATGAGCTAGCCCAACATGGGAAGATGCCAGATCAGCTGGAGTGGCTACC 975
DB 848 TTCTGTCAACAGGGTTTCAAGCGCTGCTGGCAATCTCCAAATCAGCAGGCGTGGCCACC 907

QY 976 TTTGTTCTCCAAGAAATTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAG---CT 1032
DB 908 TTTGTCATCCAGGAGGAGTTTCGACCGCTTCACTGGGTGCTGGTGTGCTCCCGCCCTCT 967

QY 1032 GAAACAACTCCAGTGGTGAATAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCT 1092
DB 968 TGGGAAGGCTCCGAAGGTTCTCAAGCGCTGCGCATCTATATAGAGAAATGAGCAGTCT 1027

QY 1092 GAGGTGGAATTTATTCATGTTTACATCCCTATGTTTGGAAACAAGAGGAGGCGAGTTCATTC 1152
DB 1028 GAAGTGGAGGTCATTCATGTGCTCCCGCCCTGGAGGAGGAGGAACGAGACTCTTAC 1087

QY 1152 CATTATCTCAAAACAGTACAGCAATCTTAAGTCACTTTTAAAGATGTGCAAGAAATATG 1212
DB 1088 CGCTACCCAGCAGCAGGCGCAAGAACCCCAAGATTTGCCCTGAAGCTGGCTCAGCTCCAG 1147

QY 1212 ATTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAATTAATCACTTTTGAG 1272
DB 1148 ACGGACCATCAGGCGAAATTCGTCAAGCTGCGAGNAGGAATGTTGTCAGCCATTTCAG 1207

QY 1272 ATTCTATTGAAGGAGTTGAATATATTGTCAGAGCTGGATGAGTCTCTCAGGGAATAAT 1332
DB 1208 TCCCTTTTCCCCAAGAGTGGAGTACATCGCCCGGCTGGTGGACACGGGACGCAATAT 1267

Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-resesgc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissue was provided by William A. Held, Roswell Park Cancer Institute, Department of Molecular and Cellular Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully acknowledge.

acknowledge:
Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://genome.ucsc.edu>

FEATURES

source

1. .3376

1. .3376

I. :3376
/organism="Mus musculus"

```
/organism="MUS MUSCULUS DOMESTICUS"
/mol_type="mRNA"
```

```
/mol_type="mrna"  
/strain="C57BL/6J"
```

```
/strain="C5/BL/6J"  
/db yref="FANTOM DB:C730003H19"
```

```
/db_xref="FANTOM_DB:C7"
/db_xref="taxon:10090"
```

```
/db_xref="taxon:10090"
/align="C7300031110"
```

```
/clone="C730003H19"
```

```
/sex="male"
```

```
/cell_type="tumor"
```

```
/tissue_type="liver"
```

```
/clone_lib:
```

```
/dev_stage:
```

235. 2824

```

/note="DIPF

```

ORIGIN

Query Match	27.9%	Score	871.2	DB	4	Length	3376
Best Local Similarity	59.8%	Pred.	No. 6.9e-230				
Matches 1513	Conservative	0	Mismatches 1008	Indels	7	Gaps	3
Qy	316	TTTTATGTTGAGCGGTATTCTCTGGAGTCACGCTTAAAAAGCTGTTGCCGATACACAGAAA	375				
Db	301	TTCTGTGTGCAGAAGCACTCTGTGGATGGGCTGCTAGCATTTATCCAGCGCATGTCGCAAG	360				
Qy	376	TATCATGGCTACATGATGGCTTAAGCGCACCATGATTTTCATGTTTGTGAAGAGGAATGAT	435				
Db	361	TCCTCGGGCTCATTTGTTCAGCAGAGCCCCCAGCACTTCCAGTTTGTGCAGAAAGCCTGAC	420				
Qy	436	CCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCCATGCTCTGGTGAGAACAGAAA	495				
Db	421	GAGTCTGGCCCCCACTCTCACCGTCTCTATTACCTGGAAATGCTTACGGCAGCGTGAG	480				
Qy	496	AATACACTGTTTTATTCTGAAATTTCCAAAACTATCATAGACGACGAGTCTTAAATGCTC	555				
Db	481	AACTCCCTCCTTACTTCGCGAGATCCCAAGAAATGTCGGAAGAGGCCCTGCTGCTGTG	540				
Qy	556	TCTTGGAGCGCTCTTTTTCGATCTTTTTTCAGGCACACTGGACTATGGAATGATTTCTCGA	615				
Db	541	TCTTGAAGCAGATCTGGACCACTTCAGGCCACACCCACCAATGGTGTCTACTCCCGA	600				
Qy	616	GAAGAAGAACTA-TTAAAGAGAAAGAAAACGCAATGGAAACAGTCGGAAATGCTTCTTACGA	674				
Db	601	GAGGAGACTACCTTGGGGAGCGCAAGCGCTGGGCGTCTTCGGAATCACTCTTATGA	660				
Qy	675	TTATCAACCAAGGAAGTGAAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATTCACGTAAA	734				
Db	661	CTTCCACAGTGAGAGCGGCTCTTCTCTTCCAGGCCAGCAATAGCCTGTTCCACTGCAG	720				
Qy	735	AGATGAGGGCCCAAGGAATTACGCACAACTTTTAAGGCCCAATCTAGTGGAAACTAG	794				
Db	721	GGATGTGTGCAAGAATGGCTTTATGTGTGCCCGATGAAGCACTGGAGACAAGACTCA	780				

Db	1861	GCTGGCTTCTCCACAGCTGCTCCATGAGCCAGAGCTTCGACATGTTGGTGAGTCACTA	1920
Qy	1929	TAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGA	1988
Db	1921	CAGCAGTGTGAGCAGCCACCTGTGTACATGTGTACAAGCTGAGCGGCCCGATGATGA	1980
Qy	1989	CCCACTTGTCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCC	2048
Db	1981	CCCACTGCACAAGCAACCCAGCTCTCTGGCCAGCATGATGGAGGAGCAATTTGCCCC	2040
Qy	2049	TGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGG	2108
Db	2041	AGCATGTGTGCCCTGAGATCTTCCACTTCCACCCGTGCAGAGCTGCAGCTCTACGG	2100
Qy	2109	GATGCTTACAGGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGCTCTTCTCAT	2168
Db	2101	CATGATCTACAAAGCACACACCTTGCACCTTGGAGAGCACCACCTGCTCTTGT	2160
Qy	2169	ATATCGTGTCTCTCAGGTGCAAGTTGGTGAATATTCGGTTTAAAGGAGTCAAGTATTCGG	2228
Db	2161	CTATGGGGCCACAGGTGCAATTTGGTGAACAACTCTCTTAAAGGCATCAATATCTGGC	2220
Qy	2229	CTTGAATACCTAGCTCTCTAGATTATGTGGTTGTAGTGATAGACAAACAGGGGATCTCG	2288
Db	2221	GCTAAATACATGGGCATCTTGGGCTATGCTGTGGTGTGATCGATGGTGGGGCTCTCG	2280
Qy	2289	TCACCGAGGGCTTAAATTTGAAGGGCGCTTTTAAATATAAATGGGTCAAATAGAAATGA	2348
Db	2281	TCAGCGGGGCTTGCACCTTCGAGGGGGCCCTGAAATATCAAATGGCGAGGTGAGATGA	2340
Qy	2349	CGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCG	2408
Db	2341	GGACCAAGTGTGAAGCTTTCAGTAGTACGTGGCTGAGAAGTATGGCTTCAATGACTGAGCG	2400
Qy	2409	TGTGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGAG	2468
Db	2401	AGTCGCCATCCATGCTGGTCTCTACGGGGCTCTCTCTCACTATGGGGCTCATCCACAA	2460
Qy	2469	GTCAATATCTTCAGGTTGCTATGCTGGGGCCAGTCACTCTGATCTTCTATGA	2528
Db	2461	GCACAAAGTGTTCAGGTTAGCCATTCGGGGCGCTCTCTCTCTCTGTGTGGATGCTTGA	2520
Qy	2529	TACAGGATACACGGAACTTATATGGGTACCTTGACCCATGACAGAAATGACAGGGCTATTCT	2588
Db	2521	CACAGGGTACACGGAACTATACATGATGTCCCGAAATATACAGCAAGGCTATGAGGC	2580
Qy	2589	AGGATCTGTGGCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATGTTTACTGCTCTT	2648
Db	2581	AGGGTCTGTAGCCCTGCATGTGGAGAAGCTGCGCAATGAGCCTAACCGCCTGTTATCTCT	2640
Qy	2649	ACATGTTTCTGGATGAGAAATGTCATTTTGCACATACCAAGTATATTTACTGAGTTTCT	2708
Db	2641	CCACGGCTTCTTGGACGAAAGCTTCACTTCTTCCACACAAATTTCTGTTGTCCTCAGCT	2700
Qy	2709	AGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAG	2768
Db	2701	GATCCGAGCAGGAAGCCATACACGCTTCAGATCTTACCCAAACAGAGACATAGCATCCG	2760
Qy	2769	AGTTCTGATCGGAGACATTTGAACTGATCTTTTGGTCACTTACCTTCAAGAAACCT	2828
Db	2761	CTGCCGAGTCCGAGAGCATTTACAGGTGACGCTGCTGACCTTTCTGAGGAACACCT	2820
Qy	2829	TGATATC	2836
Db	2821	GTGACCTC	2828
RESULT 15			
BQ675006			
LOCUS			
DEFINITION			
BQ675006 957 bp mRNA linear EST 15-JUL-2002			
AGENCOURT_8303592 NIH_MGC_102 Homo sapiens cdna clone IMAGE:6274622			
5', mRNA sequence.			
ACCESSION BQ675006			
VERSION BQ675006.1			
GI:21785840			
EST.			
KEYWORDS Homo sapiens (human)			
SOURCE			
ORGANISM Homo sapiens			
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE 1 (bases 1 to 957)			
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.			
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL Unpublished (1999)			
COMMENT Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-remail.nih.gov			
Tissue procurement: ATCC			
cdna Library Preparation: Rubin Laboratory			
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLCM2456 row: h column: 15			
High quality sequence start: 4			
High quality sequence stop: 624.			
FEATURES			
source			
1..957			
Location/Qualifiers			
1..957			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:6274622"			
/tissue_type="epidermoid carcinoma, cell line"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH_MGC_102"			
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;			
Site_2: EcoRI; cdna made by oligo-dT priming.			
Directionally cloned into EcoRI/XhoI sites using the			
following 5' adaptor: GGACGAG(G). Library constructed			
by Ling Hong in the laboratory of Gerald M. Rubin			
(University of California, Berkeley) using ZAP-cdna			
synthesis kit (Stratagene) and Superscript II RT (Life			
Technologies). Note: this is a NIH_MGC Library."			
ORIGIN			
Query Match 27.2%; Score 848.6; DB 5; Length 957;			
Best Local Similarity 96.9%; Pred. No. 8.3e-224;			
Matches 908; Conservative 0; Mismatches 24; Indels 5; Gaps 4;			
Qy	1630	AGCAATATATAACGATCCAGTGTGGGCTGCTGCTCAAGTGATTTCAAGTGTCTATC	1689
Db	9	AGGAATATAACGATCCAGTGTGGGCTGCTGCTCAAGTGATTTCAAGTGTCTATC	68
Qy	1690	AAAGGAGNTAGCAATACCAAGTGTGAATGGGAAGTTCTTGGCCGCGATGATCTAAT	1749
Db	69	AAAGGAGNTAGCAATACCAAGTGTGAATGGGAAGTTCTTGGCCGCGATGATCTAAT	128
Qy	1750	ATCAAGTTGATGAAGTCAAGAGGCTGTATATTTGAAGCACCACCAAGACTCCCTTTA	1809
Db	129	ATCAAGTTGATGAAGTCAAGAGGCTGTATATTTGAAGCACCACCAAGACTCCCTTTA	188
Qy	1810	GAGCATCACTGTACGTAGTACGTAAATCTCTGGAGAGTCAAGGCTGACTGAC	1869
Db	189	GAGCATCACTGTACGTAGTACGTAAATCTCTGGAGAGTCAAGGCTGACTGAC	248
Qy	1870	CGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTGACTTTTATAGTAAGTAT	1929
Db	249	CGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTGACTTTTATAGTAAGTAT	308
Qy	1930	AGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAGATGAC	1989
Db	309	AGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAGATGAC	368
Qy	1990	CCCACTTGCATAAACAAGGAATTTGGGCCACCATTTTGGTTCAGCAGGTCCTCTTCT	2049

Db 369 CCAACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCCT 428
Qy 2050 GACTATACCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGG 2109
Db 429 GACTATACCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGG 488
Qy 2110 ATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTATA 2169
Db 489 ATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTATA 548
Qy 2170 TATGTTGCTCTCAGGTGCGATGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGC 2229
Db 549 TATGTTGCTCTCAGGTGCGATGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGC 608
Qy 2230 TTGAATA-CCCTAGCCTCTCTAGGTTATGTGTTGTAGTGATAGAC-AACAGGGGATCCT 2287
Db 609 TTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTAGTGATAGACAAACAGGGGATCCT 668
Qy 2288 GTCAACCGAGGCTTAAATTTGAAGGCGCC--TTTAAATATAAAATGGGTCAAATAGAAAT 2345
Db 669 GTCAACCGAGGCTTAAATTTGAAGGCGCCCTTTTAAATATAAAATGGGTCAAATAGAAAT 728
Qy 2346 TGACGATCAGGTGGAAGACTCCAAATCTAGCTTCTCGATATGATTTTCAATTGACTTAGA 2405
Db 729 TGACGATCAGGGGGAAGACTCCAAATCTAGCTTCTCGATATGATTTTCAATTGACTTAGA 788
Qy 2406 TCGTGTGGGCATCCACGGCTGTCCTATGGAGGATACCTCTCCCTGATGGCATTAATGCA 2465
Db 789 TCGTGTGGGCATCCACGGCTGTCCTATGGAGGATCCCTCTCCCTGAAGGCATTTATGCA 848
Qy 2466 GAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCAGTCACCTCTG-TGGATCTTCT 2524
Db 849 GAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCAGTCACCTCTG-TGGATCTTCT 2561
Qy 2525 ATGATACAGGATACACCGAAGCTTATATGGGTCACCC 2561
Db 909 ATGATACAGGATACCCGGAACGGTATTATGGGGCCCC 945

Search completed: April 15, 2006, 00:36:30
Job time : 10968 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 19:47:01 ; Search time 533 Seconds
(without alignments)
10405.239 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgctaaagctccgagg.....agaattactataaaaaaaa 3120

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents NA.*
- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
 - 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
 - 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
 - 6: /cgn2_6/ptodata/1/ina/PCTRUS COMB.seq.*
 - 7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
 - 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
 - 9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3120	100.0	3120	3	US-10-070-464-2
2	3106.4	99.6	4829	3	US-09-976-674-12
3	2824.4	90.5	4685	3	US-09-976-674-22
4	2806.4	89.9	4676	3	US-09-976-674-20
5	2759.4	88.4	2797	3	US-09-976-594-1103
6	2644.4	84.8	2671	3	US-09-976-674-2
7	2500.4	80.1	4523	3	US-09-976-674-8
8	2056.4	65.9	4309	3	US-09-976-674-14
9	1347	43.2	1669	3	US-10-070-464-6
10	1036	33.2	1197	3	US-10-070-464-4
11	926.4	29.7	1356	3	US-09-976-674-10
12	861	27.6	2617	3	US-09-976-674-4
13	861	27.6	4219	3	US-09-976-674-28
14	861	27.6	4302	3	US-09-976-674-24
15	801.4	25.7	832	3	US-09-976-674-18
16	798.6	25.6	4180	3	US-09-976-674-36
17	798.6	25.6	4263	3	US-09-976-674-34
18	791.2	25.4	1083	3	US-10-070-464-8
19	716.8	23.0	4076	3	US-09-976-674-32
20	716.8	23.0	4159	3	US-09-976-674-30
21	654.4	21.0	4037	3	US-09-976-674-40
22	654.4	21.0	4120	3	US-09-976-674-38
23	633.4	20.3	823	3	US-09-280-116-171
24	472	15.1	620	3	US-09-976-674-16

25	376.8	12.1	2411	3	US-09-976-674-26	Sequence 26, Appl
26	326.8	10.5	612	3	US-09-392-184-31	Sequence 31, Appl
c	297	9.5	308	3	US-09-621-976-1787	Sequence 1787, Ap
c	28	131	4.2	480	3	Sequence 684, App
c	29	131	4.2	480	3	Sequence 15966, A
c	30	82.6	2.6	2946	3	Sequence 9675, Ap
c	31	82.6	2.6	14555	3	Sequence 1096, Ap
32	82	2.6	89	3	US-09-902-540-1096	Sequence 1096, Ap
33	82	2.6	1023	3	US-09-513-999C-16433	Sequence 16433, A
34	79.4	2.5	3407	3	US-09-270-767-13726	Sequence 13726, A
35	79.4	2.5	3407	3	US-10-002-593-5	Sequence 5, Appli
36	79.4	2.5	3407	3	US-09-949-016-275	Sequence 275, App
37	79.4	2.5	3419	3	US-10-423-714-5	Sequence 5, Appli
38	77.8	2.5	2924	3	US-09-949-016-4579	Sequence 4579, Ap
39	77.8	2.5	2924	6	Sequence 1026, Ap	
40	74.8	2.4	936	3	PCT-US93-07923-1	Sequence 1, Appli
41	71.8	2.3	2199	3	US-09-270-767-12940	Sequence 12940, A
c	71.2	2.3	657	3	US-09-518-550-38	Sequence 38, Appl
43	65.8	2.1	2736	3	US-09-221-017B-646	Sequence 646, App
44	65.8	2.1	2815	2	US-09-949-016-2459	Sequence 2459, Ap
45	65.8	2.1	2815	2	US-08-230-491A-1	Sequence 1, Appli
					US-08-619-280A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-070-464-2
; Sequence 2, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-2

Query Match	100.0%	Score 3120;	DB 3;	Length 3120;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3120;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	AAGTGTAAAGCCCTCCGAGGCCAAGCCGCTGCTACTCCCGCCGCTGCTTTAGTCCG	60	
Db	1	AAGTGTAAAGCCCTCCGAGGCCAAGCCGCTGCTACTCCCGCCGCTGCTTTAGTCCG	60	
Qy	61	CGTTCCGCGCTGGGTTGTACCGCGCGCGCGCGAGAGCCACTGCACACGAGCCG	120	
Db	61	CGTTCCGCGCTGGGTTGTACCGCGCGCGCGCGAGAGCCACTGCACACGAGCCG	120	
Qy	121	GAGTGGAGGCGGCGAGCATGAAGCGCGCGCGCTCCATAGCGACGCTCGGAGCG	180	
Db	121	GAGTGGAGGCGGCGAGCATGAAGCGCGCGCGCTCCATAGCGACGCTCGGAGCG	180	
Qy	181	TCCGGGCGGCGCGCGGGAAGAAATGCAACATGGCAGCAATGAAACAGACAG	240	
Db	181	TCCGGGCGGCGCGCGGGAAGAAATGCAACATGGCAGCAATGAAACAGACAG	240	
Qy	241	CTGGGTGAGATATTTGAAACTGGGAGTGTAGGAGATATTTGAATCATCAGATCGG	300	

Db	241	CTGGGTGTTGAGATATTTGAAACTGGAGCTGTGAGAGAAATATTTGAATCAAGGATCGG	300
Qy	301	CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGTCGTT	360
Db	301	CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGTCGTT	360
Qy	361	GCCGATACAGAGAAATATCATGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT	420
Db	361	GCCGATACAGAGAAATATCATGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT	420
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTAACCTTGCCATGTCT	480
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTAACCTTGCCATGTCT	480
Qy	481	GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCAAAATATCAATAGAGCA	540
Db	481	GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCAAAATATCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGATGTATTTCTCGAGAGAGACTATTAAGAGAGAAAGCGCATTTGGAACAGTCGGA	660
Db	601	GGATGTATTTCTCGAGAGAGAACTATTAAGAGAGAAAGCGCATTTGGAACAGTCGGA	660
Qy	661	ATTGCTCTTACGATATACCAAGGAAGTGGACATTTCTGTTTCAAGCCGGTAGTGA	720
Db	661	ATTGCTCTTACGATATACCAAGGAAGTGGACATTTCTGTTTCAAGCCGGTAGTGA	720
Qy	721	ATTTATACGTAAGATGGAAGGCCACAAGGATTTTACGCAACACTTTTAAGGCCAAT	780
Db	721	ATTTATACGTAAGATGGAAGGCCACAAGGATTTTACGCAACACTTTTAAGGCCAAT	780
Qy	781	CTAGTGGAACTAGTTGTCACACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAACTAGTTGTCACACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACACAGAGAA	900
Db	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACACAGAGAA	900
Qy	901	GAAGGAGACTCACCTTATGTCACATGAGTAGCCAACTAGCAAGATGCAAGATGCCAGATCA	960
Db	901	GAAGGAGACTCACCTTATGTCACATGAGTAGCCAACTAGCAAGATGCAAGATGCCAGATCA	960
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Db	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Qy	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAABATTTCTAGAAATTTCTATATGAAGNA	1080
Db	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAABATTTCTAGAAATTTCTATATGAAGNA	1080
Qy	1081	AATGATGAATCTGAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140
Db	1081	AATGATGAATCTGAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140
Qy	1141	GCAGATTCATTCGGTTATCTCAAACAGGTAACAGAAATCCTAAAGTCATCTTTAAGATG	1200
Db	1141	GCAGATTCATTCGGTTATCTCAAACAGGTAACAGAAATCCTAAAGTCATCTTTAAGATG	1200
Qy	1201	TCAGAAATTAATGATTGCTGCTGAAGGAGGATCATAGATGCTATAGATAAGGAACATAAT	1260
Db	1201	TCAGAAATTAATGATTGCTGCTGAAGGAGGATCATAGATGCTATAGATAAGGAACATAAT	1260
Qy	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTGAATATATTTGCCAGAGCTGGAGCTCT	1320
Db	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTGAATATATTTGCCAGAGCTGGAGCTCT	1320
Qy	1321	GAGGAGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTG	1380
Db	1321	GAGGAGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTG	1380

Qy	1381	TTGATCTCACCTGAATATTTATCCAGTAGAAGATGATGTTATGAAAGGAGAGACTC	1440
Db	1381	TTGATCTCACCTGAATATTTATCCAGTAGAAGATGATGTTATGAAAGGAGAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATTTCTGACCGCACTAATATCTATGAAGAAACACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCTGATTTCTGACCGCACTAATATCTATGAAGAAACACAGACATC	1500
Qy	1501	TGATAAAATATCCATGACATCTTTCTGTTTTTCCCAAAGTCAAGAGGAAATTTGAG	1560
Db	1501	TGATAAAATATCCATGACATCTTTCTGTTTTTCCCAAAGTCAAGAGGAAATTTGAG	1560
Qy	1561	TTTATTTTTGCTCTCGAATGCAAAACAGGTTTCCGTCAMTTATACAAAATACATCTAT	1620
Db	1561	TTTATTTTTGCTCTCGAATGCAAAACAGGTTTCCGTCAMTTATACAAAATACATCTAT	1620
Qy	1621	TTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCTCAAAGTGATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCTCAAAGTGATTTCAAG	1680
Qy	1681	TGTCCTATCAAAAGAGGAGATAGCAATTAACAGTGGTGGTGAATGGGAAGTTCTTGGCCCGCAT	1740
Db	1681	TGTCCTATCAAAAGAGGAGATAGCAATTAACAGTGGTGGTGAATGGGAAGTTCTTGGCCCGCAT	1740
Qy	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGCTATATTTGAAGGCACCAAGAC	1800
Db	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGCTATATTTGAAGGCACCAAGAC	1800
Qy	1801	TCCCTTTTAGAGCATCACCTGTAGTAGTCAGTTTACGTAATCTCGAGAGGTGACAAGG	1860
Db	1801	TCCCTTTTAGAGCATCACCTGTAGTAGTCAGTTTACGTAATCTCGAGAGGTGACAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCATCTTCTGTCATCAGTCAGCACTGTGACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCATCTTCTGTCATCAGTCAGCACTGTGACTTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCT	1980
Db	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCT	1980
Qy	1981	GAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCCAACATTTTGGATTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCCAACATTTTGGATTCAGCAGGT	2040
Qy	2041	CCTTTCTCTGACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACA	2100
Db	2041	CCTTTCTCTGACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACA	2100
Qy	2101	TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Db	2101	TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTTGGTCTCAGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG	2220
Db	2161	CTGTTTCATATATGTTGGTCTCAGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG	2220
Qy	2221	TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTTGTTGAGTAGACACACAGG	2280
Db	2221	TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTTGTTGAGTAGACACACAGG	2280
Qy	2281	GGATCTGTCAACGAGGCTTAAATTTCAAGCGGCTTTAAATATAAAATGAGTCAAAATA	2340
Db	2281	GGATCTGTCAACGAGGCTTAAATTTCAAGCGGCTTTAAATATAAAATGAGTCAAAATA	2340
Qy	2341	GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCATATGATTTTCATTTGAC	2400
Db	2341	GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCATATGATTTTCATTTGAC	2400
Qy	2401	TTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAAGATACCTCTCCCTGATGGCAATTA	2460
Db	2401	TTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAAGATACCTCTCCCTGATGGCAATTA	2460

QY 2461 ATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCAGTCACTCTGTGGATC 2520
DB |||||
QY 2461 ATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCAGTCACTCTGTGGATC 2520
DB |||||
QY 2521 TTCTATGATACAGGATACAGGACGTTATATGGGTCACTCCCTGACAGATGACAGGCG 2580
DB |||||
QY 2521 TTCTATGATACAGGATACAGGACGTTATATGGGTCACTCCCTGACAGATGACAGGCG 2580
DB |||||
QY 2581 TATTACTTAGGATCTGTGGCCATCAAGCAGAGAAAGTTCCCTCTCTGAACCAATCGTTTA 2640
DB |||||
QY 2581 TATTACTTAGGATCTGTGGCCATCAAGCAGAGAAAGTTCCCTCTCTGAACCAATCGTTTA 2640
DB |||||
QY 2641 CTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGCATACACAGTATATACG 2700
DB |||||
QY 2641 CTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGCATACACAGTATATACG 2700
DB |||||
QY 2701 AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACAC 2760
DB |||||
QY 2701 AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACAC 2760
DB |||||
QY 2761 AGCATAAGAGTTCTCGAATCGGAGAACATTTATGAATGCTATCTTTTGCACACTCTCAA 2820
DB |||||
QY 2761 AGCATAAGAGTTCTCGAATCGGAGAACATTTATGAATGCTATCTTTTGCACACTCTCAA 2820
DB |||||
QY 2821 GAAACCTTTGGATCAGTATTTGCTCTCTAAAGTGATATAATTTTGAACCTGTGTAGAAC 2880
DB |||||
QY 2821 GAAACCTTTGGATCAGTATTTGCTCTCTAAAGTGATATAATTTTGAACCTGTGTAGAAC 2880
DB |||||
QY 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA 2940
DB |||||
QY 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA 2940
DB |||||
QY 2941 ATTGATCATCATTTTGCATCTGACCTGCAATGAACTCTCTCTGAAATAAATGTTGGTG 3000
DB |||||
QY 2941 ATTGATCATCATTTTGCATCTGACCTGCAATGAACTCTCTCTGAAATAAATGTTGGTG 3000
DB |||||
QY 3001 CCATGCGGGGTCTACGGTTTGTGGTAGTAACTCTTAACTTAACTTAACTTAACTTAACTTAA 3060
DB |||||
QY 3001 CCATGCGGGGTCTACGGTTTGTGGTAGTAACTCTTAACTTAACTTAACTTAACTTAACTTAA 3060
DB |||||
QY 3061 TCAATGATACATATTTCTGAGAGACCCAGCAATACCAATGAGATTTACTTAAATTAATTA 3120
DB |||||
QY 3061 TCAATGATACATATTTCTGAGAGACCCAGCAATACCAATGAGATTTACTTAAATTAATTA 3120
DB |||||

RESULT 2

US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match 99.6%; Score 3106.4; DB 3; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1 AAGTGCTAAAGCTCCGAGGCCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG 60

DB |||||
QY 1 AAGTGCTAAAGCTCCGAGGCCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG 60
DB |||||
QY 61 CGTTCCCGCTTGGGTTGTTCACCGCGCGCGCGCGAGGACCTTGCACACAGGACCG 120
DB |||||
QY 61 CGTTCCCGCTTGGGTTGTTCACCGCGCGCGCGCGAGGACCTTGCACACAGGACCG 120
DB |||||
QY 121 GAGTGAGGCGCGCAGCATGAGGCGCGAGGCGCGCTCCATAGCGCAGCTCGGAGCGG 180
DB |||||
QY 121 GAGTGAGGCGCGCAGCATGAGGCGCGAGGCGCGCTCCATAGCGCAGCTCGGAGCGG 180
DB |||||
QY 181 TCCGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGCAATGGAACAGAAACAGAACAG 240
DB |||||
QY 181 TCCGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGCAATGGAACAGAAACAGAACAG 240
DB |||||
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAACAGGATCGG 300
DB |||||
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAACAGGATCGG 300
DB |||||
QY 301 CCTAAATTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360
DB |||||
QY 301 CCTAAATTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360
DB |||||
QY 361 GCCGATACAGAGAAATATCATGGCTACATGATGCTAAGGCCACCAATGATTTTCATGCTTT 420
DB |||||
QY 361 GCCGATACAGAGAAATATCATGGCTACATGATGCTAAGGCCACCAATGATTTTCATGCTTT 420
DB |||||
QY 421 GTGAAGAGAAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
DB |||||
QY 421 GTGAAGAGAAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
DB |||||
QY 481 GGTGAGAACAGAGAAATACTGTTTATTTCTGAAATTCGCAAACTATCAATAGAGCA 540
DB |||||
QY 481 GGTGAGAACAGAGAAATACTGTTTATTTCTGAAATTCGCAAACTATCAATAGAGCA 540
DB |||||
QY 541 GCAGTCTTAACTGCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
DB |||||
QY 541 GCAGTCTTAACTGCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
DB |||||
QY 601 GGAATGTTTTCGAGAGAGAAACTATTAAAGAGAGAAAGAAACGCAATGGAACAGTCCGA 660
DB |||||
QY 601 GGAATGTTTTCGAGAGAGAAACTATTAAAGAGAGAAAGAAACGCAATGGAACAGTCCGA 660
DB |||||
QY 661 ATTGCTTCTTACGATTTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
DB |||||
QY 661 ATTGCTTCTTACGATTTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
DB |||||
QY 721 ATTTATACGTAAGATGAGGCGCCACAGGATTTACGCAACACTTTTAAAGSCCAAT 780
DB |||||
QY 721 ATTTATACGTAAGATGAGGCGCCACAGGATTTACGCAACACTTTTAAAGSCCAAT 780
DB |||||
QY 781 CTAGTGAAACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCCCCGCTGATCCA 840
DB |||||
QY 781 CTAGTGAAACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCCCCGCTGATCCA 840
DB |||||
QY 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACTGTAACAGAGAA 900
DB |||||
QY 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACTGTAACAGAGAA 900
DB |||||
QY 901 GAAAGAGAGACTCACTTATGTGCAATGAGCTAGCCCAATGGAAGAGATGCCAGATCA 960
DB |||||
QY 901 GAAAGAGAGACTCACTTATGTGCAATGAGCTAGCCCAATGGAAGAGATGCCAGATCA 960
DB |||||
QY 961 GCTGGAGTCTGCTACCTTTGCTTCCAGAGAAATTTGATAGATTTCTGGCTATTGGTGG 1020
DB |||||
QY 961 GCTGGAGTCTGCTACCTTTGCTTCCAGAGAAATTTGATAGATTTCTGGCTATTGGTGG 1020
DB |||||
QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTAATAATTTCTTGAATTTCTATATCAAGAA 1080
DB |||||
QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTAATAATTTCTTGAATTTCTATATCAAGAA 1080
DB |||||
QY 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTATATCCCTATGTTGGAAACAGGAGG 1140
DB |||||

Db 1081 AATGATGAACTCGAGGTGGAAATTAATTCATGTTTACATCCCTATATGTTGAAACAAGGAGG 1140
Qy 1141 GCAGATTCATTCCTGTTATCCTAATAACAGGTAACGAAATCCTAAAGTCACCTTTTAAGATG 1200
Db 1141 GCAGATTCATTCCTGTTATCCTAATAACAGGTAACGAAATCCTAAAGTCACCTTTTAAGATG 1200
Qy 1201 TCAGAAATPAATGATGTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGAACTAAAT 1260
Db 1201 TCAGAAATPAATGATGTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGAACTAAAT 1260
Qy 1261 CAACCTTTTGAGATTCCTATTTGAAGGAGTTGAAATATATTTGCCAGAGCTGGATGGAATCCT 1320
Db 1261 CAACCTTTTGAGATTCCTATTTGAAGGAGTTGAAATATATTTGCCAGAGCTGGATGGAATCCT 1320
Qy 1321 GAGGAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGAGCTCGCCCTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGAGCTCGCCCTACAGATAGTG 1380
Qy 1381 TTGATCTCACCTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGAGCTC 1440
Db 1381 TTGATCTCACCTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGAGCTC 1440
Qy 1441 ATTGAGTCAGTCCCTGATTCCTGTGACGCCACTAATTAATCTATGAAAGAACACAGACATC 1500
Db 1441 ATTGAGTCAGTCCCTGATTCCTGTGACGCCACTAATTAATCTATGAAAGAACACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTCATGTTTCCCAAGTCACGAGAGGAAATTCAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTCATGTTTCCCAAGTCACGAGAGGAAATTCAG 1560
Qy 1561 TTTATTTTGGCTCTGAAATGAAACAGGTTTCGGTCTATTTATACAAATTAACATCTAAT 1620
Db 1561 TTTATTTTGGCTCTGAAATGAAACAGGTTTCGGTCTATTTATACAAATTAACATCTAAT 1620
Qy 1621 TTTAAAGGAAACAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Db 1621 TTTAAAGGAAACAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Qy 1681 TGCTCTATCAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db 1681 TGCTCTATCAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Qy 1741 GGATCTAATATCAAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCCACCAAGAC 1800
Db 1741 GGATCTAATATCAAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCCACCAAGAC 1800
Qy 1801 TCCCTTTAGAGCATCCTCTGATAGTCAGTTCAGTAAATCCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTAGAGCATCCTCTGATAGTCAGTTCAGTAAATCCTGGAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAGACCTGCTACTTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAGACCTGCTACTTTTATA 1920
Qy 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTACAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTACAGCTATCAAGTCT 1980
Qy 1981 GAAGATGACCCAACTTGCAAAACAAAGGAAATTTGGGCCACCAATTTTGGATTCAGAGGT 2040
Db 1981 GAAGATGACCCAACTTGCAAAACAAAGGAAATTTGGGCCACCAATTTTGGATTCAGAGGT 2040
Qy 2041 CCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
Db 2041 CCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
Qy 2101 TTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTG 2160
Db 2101 TTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTG 2160
Qy 2161 CTGTTTCATATATGGTGG--TCCTCAGGTGCAATGGTGAATATTCGGTTTAAAGGAGTCA 2218
Db 2161 CTGTTTCATATATGGTGGTCTCCTCAGGTGCAATGGTGAATATTCGGTTTAAAGGAGTCA 2220

Qy 2219 AGTATTTCCGCTTCAATACCTAGACCTCTCTAGTTTATGTTGGTTAGTAGACAAACA 2278
Db 2221 AGTATTTCCGCTTCAATACCTAGACCTCTCTAGTTTATGTTGGTTAGTAGACAAACA 2280
Qy 2279 GGGATCTGTGTCACCGAGGCTTAAATTTGAAGGCCCTTTAAATATAAAATGGGTCAAA 2338
Db 2281 GGGATCTGTGTCACCGAGGCTTAAATTTGAAGGCCCTTTAAATATAAAATGGGTCAAA 2340
Qy 2339 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATGATTTCAATG 2398
Db 2341 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATGATTTCAATG 2400
Qy 2399 ACTTAGATCTGTGTGGGCATCCACGGCTGCTTATGGAAGGATACCTCTCCCTGATGGCAT 2458
Db 2401 ACTTAGATCTGTGTGGGCATCCACGGCTGCTTATGGAAGGATACCTCTCCCTGATGGCAT 2460
Qy 2459 TAAATCAGAGGTTCAGATATCTTCAAGGTTCTATTTGCTGGGGCCCAAGTCACTCTGTGA 2518
Db 2461 TAAATCAGAGGTTCAGATATCTTCAAGGTTCTATTTGCTGGGGCCCAAGTCACTCTGTGA 2520
Qy 2519 TCTTCTATGATACAGGATACACGGAACTTATATGGGTCAACCTTGACAGAAATGAACAGG 2578
Db 2521 TCTTCTATGATACAGGATACACGGAACTTATATGGGTCAACCTTGACAGAAATGAACAGG 2580
Qy 2579 GCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTT 2638
Db 2581 GCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTT 2640
Qy 2639 TACTGCTCTTACATGGTTTCTGTCATGAGAAATGTCCATTTTGACATACACAGTATATTAC 2698
Db 2641 TACTGCTCTTACATGGTTTCTGTCATGAGAAATGTCCATTTTGACATACACAGTATATTAC 2700
Qy 2699 TGAGTTTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGAC 2758
Db 2701 TGAGTTTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGAC 2760
Qy 2759 ACAGCATAGAGTTCTCGAATCGGGAGAACATTTATGAACTGTGATCTTTTGCACTACTTC 2818
Db 2761 ACAGCATAGAGTTCTCGAATCGGGAGAACATTTATGAACTGTGATCTTTTGCACTACTTC 2820
Qy 2819 AAGAAAACTTGGATTCAGTATTCCTGCTCTAAAGTGATATAATTTTGACCTGTGATA 2878
Db 2821 AAGAAAACTTGGATTCAGTATTCCTGCTCTAAAGTGATATAATTTTGACCTGTGATA 2880
Qy 2879 ACTCTCTGATPACACTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAAAACA 2938
Db 2881 ACTCTCTGATPACACTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAAAACA 2940
Qy 2939 GAATTTGATCATCACATTTTGTATCTGCAATGTAACTCTAATACCTTAAACCCCATGCTCAA 2998
Db 2941 GAATTTGATCATCACATTTTGTATCTGCAATGTAACTCTAATACCTTAAACCCCATGCTCAA 3000
Qy 2999 TGCCATGCAAGGCTCTACGGTTTGTGGTAGTAAATCTAATACCTTAAACCCCATGCTCAA 3058
Db 3001 TGCCATGCAAGGCTCTACGGTTTGTGGTAGTAAATCTAATACCTTAAACCCCATGCTCAA 3060
Qy 3059 AATCAATGATACATATTTCTTGAGAGCCAGCAATACCATAGAAATTTCTAAAAAATA 3118
Db 3061 AATCAATGATACATATTTCTTGAGAGCCAGCAATACCATAGAAATTTCTAAAAAATA 3120
Qy 3119 AA 3120
Db 3121 AA 3122

RESULT 3
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen

APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976,674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22
LENGTH: 4685
TYPE: DNA
ORGANISM: Homo sapiens
US-09-976-674-22

Query Match 90.5%; Score 2824.4; DB 3; Length 4685;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;

QY 1 AAGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTCTTCTTAGTGCGG 60
DB 1 AAGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTCTTCTTAGTGCGG 60

QY 61 CGTTCGCCGCTGGGTGTGTCACCGCGCGCGCGCGAGGAAGCACTGCAACAGGACCG 120
DB 61 CGTTCGCCGCTGGGTGTGTCACCGCGCGCGCGCGAGGAAGCACTGCAACAGGACCG 120

QY 121 GAGTGAGGCGGCGAGCATGAAGCGCGCGCGCGCGCTCCATAGCCAGCTCGGACCG 180
DB 121 GAGTGAGGCGGCGAGCATGAAGCGCGCGCGCGCGCTCCATAGCCAGCTCGGACCG 180

QY 181 TCCGGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAACACAGAACAG 240
DB 181 TCCGGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAACACAGAACAG 240

QY 241 CTGGGTGTTGAGATATTGAAACTCGGACCTGTGAGGAGATAATTGAATCACAGGATCGG 300
DB 241 CTGGGTGTTGAGATATTGAAACTCGGACCTGTGAGGAGATAATTGAATCACAGGATCGG 300

QY 301 CTTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360
DB 301 CTTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360

QY 361 GCCGATACAGAAATATCATGCTACATGATGGCTAGGCAACACATGATTTTCATGTTT 420
DB 361 GCCGATACAGAAATATCATGCTACATGATGGCTAGGCAACACATGATTTTCATGTTT 420

QY 421 GTCAAGAGGAATGATCCAGATGGAACCTCATTACAGACAGAACTCTATTACCTTGGCATGCT 480
DB 421 GTCAAGAGGAATGATCCAGATGGAACCTCATTACAGACAGAACTCTATTACCTTGGCATGCT 480

QY 481 GGTGAGAACAGAGAAATACATGTTTATCTGAAATTTCCCAAACTATCAATAGAGCA 540
DB 481 GGTGAGAACAGAGAAATACATGTTTATCTGAAATTTCCCAAACTATCAATAGAGCA 540

QY 541 GCAGTCTTAATGCTCTCTTGGAGCTCTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
DB 541 GCAGTCTTAATGCTCTCTTGGAGCTCTTTGGATCTTTTTCAGGCAACACTGGACTAT 600

QY 601 GGAATGTATTCTCGAAGAGAACTATTAAAGAGAAAGAAACGCATTGGAAACAGTCGGA 660
DB 601 GGAATGTATTCTCGAAGAGAACTATTAAAGAGAAAGAAACGCATTGGAAACAGTCGGA 660

QY 661 ATTGCTCTTTACGATTATCAACAGGAAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
DB 661 ATTGCTCTTTACGATTATCAACAGGAAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720

QY 721 ATTTATCACGTAAAGATGAGGGCCACAGGATTTACGCAACACCTTTAAGGCCCAAT 780
DB 721 ATTTATCACGTAAAGATGAGGGCCACAGGATTTACGCAACACCTTTAAGGCCCAAT 780

QY 781 CTAGTGAATACTAGTTGTCCTCCACATACGATGGATCCAAATATTATGCCCCGCTGATCCA 840

DB 781 CTAGTGAATACTAGTTGTCCTCCACATACGATGGATCCAAATATTATGCCCTGCTGATCCA 840

QY 841 GACTGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACTCGTAACACAGAGAA 900
DB 841 GACTGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACTCGTAACACAGAGAA 900

QY 901 GAAAGGAGACTCATCTTATGTGCAATGAGCTAGCAACATGGAAGAAGATGCCAGATCA 960
DB 901 GAAAGGAGACTCATCTTATGTGCAATGAGCTAGCAACATGGAAGAAGATGCCAGATCA 960

QY 961 GCTGGAGTGCCTACCTTTGTTCTCCAAGAAGAAATTTGATAGATATCTTGGCTATTGTTGG 1020
DB 961 GCTGGAGTGCCTACCTTTGTTCTCCAAGAAGAAATTTGATAGATATCTTGGCTATTGTTGG 1020

QY 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
DB 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGAA 1080

QY 1081 AATGATGAATCTGAGGTGGAATTAATTCATGTTTACATCCCTATGTTGGAAACAGGAGG 1140
DB 1081 AATGATGAATCTGAGGTGGAATTAATTCATGTTTACATCCCTATGTTGGAAACAGGAGG 1140

QY 1141 GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATG 1200
DB 1141 GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATG 1200

QY 1201 TCAGAAATTAATGATGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAATCAAT 1260
DB 1201 TCAGAAATTAATGATGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAATCAAT 1260

QY 1261 CAACCTTTTGAAGATCTTATTTGAAGAGTTGAATATATGTCAGAGCTGGATGAGTCCCT 1320
DB 1261 CAACCTTTTGAAGATCTTATTTGAAGAGTTGAATATATGTCAGAGCTGGATGAGTCCCT 1320

QY 1321 GAGGGAATAATGCTTGGTCCATCTTACTAGATCGCTCCAGACTCCGCTTACAGATAGT 1380
DB 1321 GAGGGAATAATGCTTGGTCCATCTTACTAGATCGCTCCAGACTCCGCTTACAGATAGT 1380

QY 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGACAGAGACTC 1440
DB 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGACAGAGACTC 1440

QY 1441 ATTGAGTCAAGTGCCTGATTTCTGTGACGCGCACTTAATTTATGTAAGAAACAAACAGATC 1500
DB 1441 ATTGAGTCAAGTGCCTGATTTCTGTGACGCGCACTTAATTTATGTAAGAAACAAACAGATC 1500

QY 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCAGAGAGGAATTTGAG 1560
DB 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCAGAGAGGAATTTGAG 1560

QY 1561 TTTATTTTGGCTCTCAATGCAAAACAGGTTTTCCTCATTTTATACAAAATTTACATCTATT 1620
DB 1561 TTTATTTTGGCTCTCAATGCAAAACAGGTTTTCCTCATTTTATACAAAATTTACATCTATT 1620

QY 1621 TTTAAAGGAACAAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
DB 1621 TTTAAAGGAACAAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680

QY 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGCAT 1740
DB 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGCAT 1740

QY 1741 GGATCTTAATCAAGTGGATGATGAAGTCAGAGGCTGTTATTTTGAAGGACCAACAGAC 1800
DB 1741 GGATCTTAATCAAGTGGATGATGAAGTCAGAGGCTGTTATTTTGAAGGACCAACAGAC 1800

QY 1801 TCCCTTTAGAGCATCACCTGTACGTAGTCACTAGTAAATCTCTGGAGAGGTGACAAGG 1860
DB 1801 TCCCTTTAGAGCATCACCTGTACGTAGTCACTAGTAAATCTCTGGAGAGGTGACAAGG 1860

QY 1861 CTGATCGACCGTGGCTACTCATTCTTGTCTGCTCATCTAGCTAGCTGCTCTTTTATA 1920

Db 1861 CTGACTGACCGTGCTACTCACAATCTTGCTGTCATCAGTGCAGCACTGTGACTTCTTTATA 1920
QY 1921 AGTAAGTATAGTAACCAAGAAATPCCACACTGTGTGCCCTTTACAAGCTATCAAGTCCCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATPCCACACTGTGTGCCCTTTACAAGCTATCAAGTCCCT 1980
QY 1981 GAAAGATGACCCCAACTTGCMAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT 2040
Db 1981 GAAAGATGACCCCAACTTGCMAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT 2034
QY 2041 CCTCTTCTGACTACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
Db 2035 ----- 2034
QY 2101 TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
Db 2035 ----- 2034
QY 2161 CTGTTTCATATATGTTGGTCTCAGGTGCAGTGTGTAATTCGGTTTAAAGGATCAAG 2220
Db 2035 -----GTCCCTCAGGTGCAGTGTGTAATTCGGTTTAAAGGATCAAG 2078
QY 2221 TATTTCCGGTTGAATACCTTAGCCTCTAGGTTATGTGGTTAGTGTAGTATAGACACAGG 2280
Db 2079 TATTTCCGGTTGAATACCTTAGCCTCTAGGTTATGTGGTTAGTGTAGTATAGACACAGG 2138
QY 2281 GGATCCTGTCCCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATATGGGTCAATA 2340
Db 2139 GGATCCTGTCCCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATATGGGTCAATA 2198
QY 2341 GAAATTTGAGGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGAC 2400
Db 2199 GAAATTTGAGGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGAC 2258
QY 2401 TTAGATCGTGTGGGATCCAGGGTGTCTTATGGAGGATACCTCTCCCTGTATGGCATTA 2460
Db 2259 TTAGATCGTGTGGGATCCAGGGTGTCTTATGGAGGATACCTCTCCCTGTATGGCATTA 2318
QY 2461 ATGCAGGTCAGATATCTTCAGGGTGTCTTATGGGGCCCGAGTCACTCTGTGGATC 2520
Db 2319 ATGCAGGTCAGATATCTTCAGGGTGTCTTATGGGGCCCGAGTCACTCTGTGGATC 2378
QY 2521 TTCTATGATACAGGATACACGGAGTTATATGGTTCACCTTGACAGATGAACAGGCG 2580
Db 2379 TTCTATGATACAGGATACACGGAGTTATATGGTTCACCTTGACAGATGAACAGGCG 2438
QY 2581 TATTACTTATAGGATCTGTGGCCATGCAAGCAAGATTCGCCCTTGAAACCAATCGTTTA 2640
Db 2439 TATTACTTATAGGATCTGTGGCCATGCAAGCAAGATTCGCCCTTGAAACCAATCGTTTA 2498
QY 2641 CTGCTCTTACATGGTTTCTCGATGAGAAATGTCCATTTTGACATACCACTATATCTG 2700
Db 2499 CTGCTCTTACATGGTTTCTCGATGAGAAATGTCCATTTTGACATACCACTATATCTG 2558
QY 2701 AGTTTTTTATGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACAC 2760
Db 2559 AGTTTTTTATGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACAC 2618
QY 2761 AGCATAGAGTTCTCGAATCGGGAGAACATTTGAACATGCAATCTTTTGGCACTACCTTCAA 2820
Db 2619 AGCATAGAGTTCTCGAATCGGGAGAACATTTGAACATGCAATCTTTTGGCACTACCTTCAA 2678
QY 2821 GAAAACTTGGATCAGTATGCTGCTCTTAAAGTGATATATTTTGAACCTGTGTAGAAC 2880
Db 2679 GAAAACTTGGATCAGTATGCTGCTCTTAAAGTGATATATTTTGAACCTGTGTAGAAC 2738
QY 2881 TCTCTGTGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGA 2940
Db 2739 TCTCTGTGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGA 2798
QY 2941 ATTGATCATCATAATTTTGTATCTGCAATGTAACATCTACTCTGAAAAATAAATGTGGTG 3000
Db 2799 ATTGATCATCATAATTTTGTATCTGCAATGTAACATCTACTCTGAAAAATAAATGTGGTG 2858

QY 3001 CCATCAGGCGCTACGGTTTGGTAGTAATCTTAATACCTTAAACCCACATGCTCAAAA 3060
Db 2859 CCATCAGGCGCTACGGTTTGGTAGTAATCTTAATACCTTAAACCCACATGCTCAAAA 2918
QY 3061 TCAATATGATACATATTTCTGAGAGCCAGCAGCATACATAGAAATCTTAAAAAAGAAAA 3120
Db 2919 TCAATATGATACATATTTCTGAGAGCCAGCAGCATACATAGAAATCTTAAAAAAGAAAA 2978

RESULT 4
US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-976-674-20

Query Match 89.9%; Score 2806.4; DB 3; Length 4676;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;

QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGCTGTCTACTGCCCGCGCTCTTTAGTCCG 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGCTGTCTACTGCCCGCGCTCTTTAGTCCG 60

QY 61 CGTTCCCGCCCTGGGTTGTCAACCGCGCGCGCGCGAGGAAGCACTGCAACACAGACCG 120
Db 61 CGTTCCCGCCCTGGGTTGTCAACCGCGCGCGCGCGAGGAAGCACTGCAACACAGACCG 120

QY 121 GAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGG 180
Db 121 GAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGG 180

QY 181 TCCGGCGCGCGCGCGGGGAGGAAAAATGCAACATGCGAGCGCGCGCTCCATAGCGCACAG 240
Db 181 TCCGGCGCGCGCGCGGGGAGGAAAAATGCAACATGCGAGCGCGCGCTCCATAGCGCACAG 240

QY 241 CTGGGTGTTGAGATATTTGAAACTCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300

QY 301 CCTAAATTCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
Db 301 CCTAAATTCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360

QY 361 GCCGATACCAAGAAATATCATGCTACATGCTAGGCTAGGCAACACATGATTTTCATGCTT 420
Db 361 GCCGATACCAAGAAATATCATGCTACATGCTAGGCTAGGCAACACATGATTTTCATGCTT 420

QY 421 GTGAAGAGGAATGATTCAGATGAGCTCATTTTCAGACAGAACTTATACCTTGCATCTCT 480
Db 421 GTGAAGAGGAATGATTCAGATGAGCTCATTTTCAGACAGAACTTATACCTTGCATCTCT 480

QY 481 GGTGAGAACAGAGAAAAATACACTGTGTTTATTTCTGAAATTCGAAAACTATCAATAGACA 540
Db 481 GGTGAGAACAGAGAAAAATACACTGTGTTTATTTCTGAAATTCGAAAACTATCAATAGACA 540

Qy 541 GCAATCTTAATGCTCTCTTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGATTTCTCGAGAAAGAACTATTAAGAGAAAGAAAGCAATTTGCGGCAAGTCGGA 660
Db 601 GGAATGATTTCTCGAGAAAGAACTATTAAGAGAAAGAAAGCAATTTGCGGCAAGTCGGA 660
Qy 661 ATTGCTTTCTTACGATTTATCAACGAAGAGTGGAAATTTCTGTGTTTCAAGCCGGTAGTGA 720
Db 661 ATTGCTTTCTTACGATTTATCAACGAAGAGTGGAAATTTCTGTGTTTCAAGCCGGTAGTGA 720
Qy 721 ATTATACAGTAAAGATGGAGGGCCACAAGATTTAGCGAAACCACTTTAAGGCCCAAT 780
Db 721 ATTATACAGTAAAGATGGAGGGCCACAAGATTTAGCGAAACCACTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAAACTAGTTGTCTCCCAACATACGGATGGATCCAAATTTATGCCCTGTGATCCA 840
Db 781 CTAGTGGAAACTAGTTGTCTCCCAACATACGGATGGATCCAAATTTATGCCCTGTGATCCA 840
Qy 841 GACTGGATTCCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA 900
Db 841 GACTGGATTCCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAAGATGCCAGATCA 960
Qy 961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
Qy 1021 TGTCCTAAAGCTGAACAACCTCCAGTGTGTGTAATTTCTAGAAATTTCTATATGAAGAA 1080
Db 1021 TGTCCTAAAGCTGAACAACCTCCAGTGTGTGTAATTTCTAGAAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTCTATGTTGGAAACAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTCTATGTTGGAAACAGGAGG 1140
Qy 1141 GCAGATTCATTCGGTTATCTTAAACAGGTAACAGAAATCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCATTCGGTTATCTTAAACAGGTAACAGAAATCTTAAAGTCACTTTTAAAGATG 1200
Qy 1201 TCAGAAATATGATTTGATGCTGAAGGAGATCAGATGTCATAGATAAGGAACTAAT 1260
Db 1201 TCAGAAATATGATTTGATGCTGAAGGAGATCAGATGTCATAGATAAGGAACTAAT 1260
Qy 1261 CAACCTTTTGAGATTTCTATTGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCT 1320
Db 1261 CAACCTTTTGAGATTTCTATTGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCT 1320
Qy 1321 GAGGAAATATGCTTGGTCCATCTTACATGATTCGCTCCAGACTCGCTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTGGTCCATCTTACATGATTCGCTCCAGACTCGCTACAGATAGTG 1380
Qy 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTGGAAGGACAGAGACTC 1440
Db 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTGGAAGGACAGAGACTC 1440
Qy 1441 ATTGAGTCAGTGGCTGATTTCTGTGAGCGCACTAAATTTATGATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTGGCTGATTTCTGTGAGCGCACTAAATTTATGATGAAGAAACACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTCCTCCAAAGTCACGAAGAGGAAATGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTCCTCCAAAGTCACGAAGAGGAAATGAG 1560
Qy 1561 TTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATT 1620
Qy 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAGTGATTTCAAG 1680

Db 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAGTGATTTCAAG 1680
Qy 1681 TGTCTCTATCAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCCTTGGCCGCAT 1740
Db 1681 TGTCTCTATCAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCCTTGGCCGCAT 1740
Qy 1741 GGAATCAATATCAAGTTCATGAAGTCAAGAGGCTGGTATATTTTGAAGGACCAAGAGAC 1800
Db 1741 GGAATCAATATCAAGTTCATGAAGTCAAGAGGCTGGTATATTTTGAAGGACCAAGAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTACGTAGTAAATCTCGAGAGGTGACAAAG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTACGTAGTAAATCTCGAGAGGTGACAAAG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCACTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCACTGTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCCT 1980
Qy 1981 GAAGATGACCCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT 2040
Qy 2041 CCTCTTCTGACTATATCTCTCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTTACA 2100
Db 2041 CCTCTTCTGACTATATCTCTCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTTACA 2100
Qy 2101 TTGTATGGGATGCTCTTACAAAGCTCATGATCTACAGCTCGGAAGAAATATCTTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTTACAAAGCTCATGATCTACAGCTCGGAAGAAATATCTTACTGTG 2160
Qy 2161 CTGTTTCAATATGCTGGTCTCTCAGGTGCAATGTTGTTGTAATTCGTTTTAAGGAGTCAAG 2220
Db 2161 CTGTTTCAATATGCTGGTCTCTCAGGTGCAATGTTGTTGTAATTCGTTTTAAGGAGTCAAG 2220
Qy 2221 TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTATAGACACAGG 2280
Db 2281 TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTATAGACACAGG 2280
Qy 2281 GGAATTCGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGTCAATA 2340
Db 2281 GGAATTCGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGTCAATA 2340
Qy 2341 GAAATTCAGCATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2341 GAAATTCAGCATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Qy 2401 TTAGATCTGTGGGCTATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
Db 2401 TTAGATCTGTGGGCTATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
Qy 2461 ATGCAGAGGTGAGATATCTTTCAGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATC 2520
Db 2461 ATGCAGAGGTGAGATATCTTTCAGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATC 2520
Qy 2521 TTCTATGATACAGGATACAGGAACTGTTATGTTGGTTCACCTGACAGAAATGACAGGGC 2580
Db 2521 TTCTATGATACAGGATACAGGAACTGTTATGTTGGTTCACCTGACAGAAATGACAGGGC 2580
Qy 2581 TATTACTTAGGATCTGGCCATGCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTA 2640
Db 2581 TATTACTTAGGATCTGGCCATGCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTA 2640
Qy 2641 CTGCTCTTACATGTTTCTCTGGATGAGATGTCATTTTGGACATACCAAGTATATCTG 2700
Db 2641 CTGCTCTTACATGTTTCTCTGGATGAGATGTCATTTTGGACATACCAAGTATATCTG 2700
Qy 2701 AGTTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC 2760
Db 2701 AGTTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC 2760

Db	2550	AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCCTCAGAGAGACAC	2609	Db	241	ATCTTTTTCAGGCAACACTGGA	CTAGTAATGTAATTTCTCGAGAAGAAGAACTATTAAAGAG	300		
Qy	2761	AGCATAAGAGTTCCTGGAATCGGAGAAACATTTATGAATCCATCTTTTGGACATACCTTCAA	2820	Qy	635	AAAGAAAACGATTTGGAACAGT	CGGAATTCCTTACGATTATCACC	AAAGAACTGGAA	694	
Db	2610	AGCATAAGAGTTCCTGGAATCGGAGAAACATTTATGAATCCATCTTTTGGACATACCTTCAA	2669	Db	301	AAAGAAAACGATTTGGAACAGT	CGGAATTCCTTACGATTATCACC	AAAGAACTGGAA	360	
Qy	2821	GAAGAACCTTTGGATCACGATTTGCTGCTCTAAAGAGTGATATAAATTTTGACCTGTGTAGAAC	2880	Qy	695	CATTTCTGTTTCAAGCCGGTAGTGGAA	TTTATCAGCTTAAGATGAGGGCCACAGGAT	754		
Db	2670	GAAGAACCTTTGGATCACGATTTGCTGCTCTAAAGAGTGATATAAATTTTGACCTGTGTAGAAC	2729	Db	361	CATTTCTGTTTCAAGCCGGTAGTGGAA	TTTATCAGCTTAAGATGAGGGCCACAGGAT	420		
Qy	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAGAAAACACAGA	2940	Qy	755	TTACGCAACAAACCTTTTAAGGCCCAAT	CTAGTGAAATCTAGTGGAACTAGTTGTC	TCCCAACATACCGATGG	814	
Db	2730	TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAGAAAACACAGA	2789	Db	421	TTACGCAACAAACCTTTTAAGGCCCAAT	CTAGTGAAATCTAGTGGAACTAGTTGTC	TCCCAACATACCGATGG	480	
Qy	2941	ATTGATCATCACATTTTGGATACCTGCCATGTAACATCTACTCTCTGAAAATAAATGTGGTG	3000	Qy	815	ATCCAAAATATATGCCCCGCTGAT	CCAGACTGGATTCCTTTTATACATAGCAACGATATTT	874		
Db	2790	ATTGATCATCACATTTTGGATACCTGCCATGTAACATCTACTCTCTGAAAATAAATGTGGTG	2849	Db	481	ATCCAAAATATATGCCCCGCTGAT	CCAGACTGGATTCCTTTTATACATAGCAACGATATTT	540		
Qy	3001	CCATCGAGGGTCTACGGTTTGGTAGTAATCTTAATACCTTTAAACCCACATGCTCAAAA	3060	Qy	875	GGATATCTAAACATCGTAAC	CCAGAGAGAAAGGAGACTCACTTATGTGCA	CAATGAGCTAG	934	
Db	2850	CCATCGAGGGTCTACGGTTTGGTAGTAATCTTAATACCTTTAAACCCACATGCTCAAAA	2909	Db	541	GGATATCTAAACATCGTAAC	CCAGAGAGAAAGGAGACTCACTTATGTGCA	CAATGAGCTAG	600	
Qy	3061	TCAATATGATACATATCTCTGAGACCCAGCAATACATAGATTAATCTTAAAGAAAATAA	3120	Qy	935	CCAAATCGAAAGAGATGCCAGAT	CAGCTGGAGTCCGCTACCTTTGTTCTC	CCAAAGAAAGAT	994	
Db	2910	TCAATATGATACATATCTCTGAGACCCAGCAATACATAGATTAATCTTAAAGAAAATAA	2969	Db	601	CCAAATCGAAAGAGATGCCAGAT	CAGCTGGAGTCCGCTACCTTTGTTCTC	CCAAAGAAAGAT	660	
RESULT 5					TTGATAGATATTTCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTA					1054
US-09-976-594-1103					TTGATAGATATTTCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTA					720
; Sequence 1103, Application US/09976594					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1114
; Patent No. 6673549					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					780
; GENERAL INFORMATION:					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1115
; APPLICANT: Buchbinder, Jenny					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1173
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					840
; FILE REFERENCE: PA-0041 US					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1233
; CURRENT APPLICATION NUMBER: US/09/976,594					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					900
; CURRENT FILING DATE: 2001-10-12					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1293
; PRIOR APPLICATION NUMBER: 60/240,409					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					960
; PRIOR FILING DATE: 2000-10-12					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1349
; NUMBER OF SEQ ID NOS: 1143					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1020
; SOFTWARE: PERL Program					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1409
; SEQ ID NO 1103					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1080
; LENGTH: 2797					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1469
; TYPE: DNA					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1140
; ORGANISM: Homo sapiens					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1529
; FEATURE:					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1200
; NAME/KEY: misc feature					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1141
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1589
US-09-976-594-1103					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1260
Query Match 88.4%; Score 2759.4; DB 3; Length 2797;					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1590
Best Local Similarity 99.8%; Pred. No. 0;					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1649
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;					AAATTCCTTAGAATTCCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTTA					1320
Qy	335	CCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACAGAAAATATCATGGCTACATGATGG	394	Qy	1294	TATATTGCCAGAGCTGGATGGACT	CCTGAGGGAAAAATAGTGCCTTGGTCCATCCTACT	1349		
Db	1	CCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACAGAAAATATCATGGCTACATGATGG	60	Db	961	TATATTGCCAGAGCTGGATGGACT	CCTGAGGGAAAAATAGTGCCTTGGTCCATCCTACT	1020		
Qy	395	CTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG	454	Qy	1350	AGATCGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCTGAAATTTATTTATCCCAAGT	1409			
Db	61	CTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG	120	Db	1021	AGATCGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCTGAAATTTATTTATCCCAAGT	1080			
Qy	455	ACAGAAATCTATTACCTTGCATGCTCTGGTGAGAAACAGAGAAAATACACTGTTTATTCG	514	Qy	1410	AGAAATGATGTTTATGGAAGGCGAGAGACT	CATTGAGTCACTGCTGCTGATTTCTGTGACGCC	1469		
Db	121	ACAGAAATCTATTACCTTGCATGCTCTGGTGAGAAACAGAGAAAATACACTGTTTATTCG	180	Db	1081	AGAAATGATGTTTATGGAAGGCGAGAGACT	CATTGAGTCACTGCTGCTGATTTCTGTGACGCC	1140		
Qy	515	AAATTCCTCAAACTATCAATAGACGAGCTCTTAATGCTCTCTTGGAGCCCTCTTTTGG	574	Qy	1470	ACTTAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTATG	1529			
Db	181	AAATTCCTCAAACTATCAATAGACGAGCTCTTAATGCTCTCTTGGAGCCCTCTTTTGG	240	Db	1141	ACTTAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTATG	1200			
Qy	575	ATCTTTTTCAGGCAACACTGGAATGTAATTTCTCGAAGAAGAACTATTAAAGAG	634	Qy	1530	TTTTTCCCAAGCTCAGAGAGGAATTTGAGTTTATTTTGGCTCTCTCAATGCAAAACAGG	1589			
					TTTTTCCCAAGCTCAGAGAGGAATTTGAGTTTATTTTGGCTCTCTCAATGCAAAACAGG					1260

```
QY 1710 CAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1769
Db 1381 CAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1440
QY 1770 AAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACTGTAGCTAGT 1829
Db 1441 AAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACTGTAGCTAGT 1500
QY 1830 CAGTTACGTAAATCTGAGAGGTGACAAAGGCTGACTGACCGGTGCTACTCAATTTCTTG 1889
Db 1501 CAGTTACGTAAATCTGAGAGGTGACAAAGGCTGACTGACCGGTGCTACTCAATTTCTTG 1560
QY 1890 CTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGATCCACA 1949
Db 1561 CTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGATCCACA 1620
QY 1950 CTGTGTGCTCTTTTACAGCTATCAAGTCTGTAGATGACCAACTTGCARAAACAAGGA 2009
Db 1621 CTGTGTGCTCTTTTACAGCTATCAAGTCTGTAGATGACCAACTTGCARAAACAAGGA 1680
QY 2010 ATTTTGGGCGACATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCCTCCAGAAAT 2069
Db 1681 ATTTTGGGCGACATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCCTCCAGAAAT 1740
QY 2070 TTTCTCTTTTGAAGTACTACTCGGATTTACATTTGATGGATGCTCTACAAGCCTCATGA 2129
Db 1741 TTTCTCTTTTGAAGTACTACTCGGATTTACATTTGATGGATGCTCTACAAGCCTCATGA 1800
QY 2130 TCTACAGCTGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTGGTCTCAGGTGCA 2189
Db 1801 TCTACAGCTGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTGGTCTCAGGTGCA 1860
QY 2190 GTTGTGTAATACGTTTAAAGAGTCAAGTATTTCCGCTTGCAATCACTGACCTCTCT 2249
Db 1861 GTTGTGTAATACGTTTAAAGAGTCAAGTATTTCCGCTTGCAATCACTGACCTCTCT 1920
QY 2250 AGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA 2309
Db 1921 AGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA 1980
QY 2310 AGGCGCTTTAAATATAAATATGAGTCAAGTAAATGAGAAATGACGATCAGGTGGAAGTCCA 2369
Db 1981 AGGCGCTTTAAATATAAATATGAGTCAAGTAAATGAGAAATGACGATCAGGTGGAAGTCCA 2040
QY 2370 ATATCTAGCTTCTCGATATGATTTTCAATGCTTAGATCGTGGGATCCACGGCTGGTC 2429
Db 2041 ATATCTAGCTTCTCGATATGATTTTCAATGCTTAGATCGTGGGATCCACGGCTGGTC 2100
QY 2430 CTATGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTGC 2489
Db 2101 CTATGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTGC 2160
QY 2490 TATTCGTGGGCCCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAACGTTA 2549
Db 2161 TATTCGTGGGCCCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAACGTTA 2220
QY 2550 TATGGGTCACTGACAGATGACAGGGCTATTACTAGGATCTGTGGCATGCAAGC 2609
Db 2221 TATGGGTCACTGACAGATGACAGGGCTATTACTAGGATCTGTGGCATGCAAGC 2280
QY 2610 AGAAAAGTTCCCTCTGTAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAA 2669
Db 2281 AGAAAAGTTCCCTCTGTAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAA 2340
QY 2670 TGTCCATTTTGCACATACCAAGTATATTAATCAGTTTTTTTGTAGTGGGCTGGAAAGCCATA 2729
Db 2341 TGTCCATTTTGCACATACCAAGTATATTAATCAGTTTTTTTGTAGTGGGCTGGAAAGCCATA 2400
QY 2730 TGAATTTACAGATCTATCTCAGGAGAGACACAGCATAAAGATTCCTGAAATCGGGAGAAC 2789
Db 2401 TGAATTTACAGATCTATCTCAGGAGAGACACAGCATAAAGATTCCTGAAATCGGGAGAAC 2460
```

```
QY 2790 TTATGAATGTCATCTTTTGGCACTACCTTCAAGAAAACTTTGGATCAGTATTCGTCTCT 2849
Db 2461 TTATGAATGTCATCTTTTGGCACTACCTTCAAGAAAACTTTGGATCAGTATTCGTCTCT 2520
QY 2850 AAAAGTGATATATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAA 2909
Db 2521 AAAAGTGATATATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAA 2580
QY 2910 ATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATATTTTGTGCTGAT 2969
Db 2581 ATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATATTTTGTGCTGAT 2640
QY 2970 GTAAACATCTACTCTGAAATAAATGTGTCCTCATGAGGGGTCTACGGTTTGTGTAGT 3029
Db 2641 GTAAACATCTACTCTGAAATAAATGTGTCCTCATGAGGGGTCTACGGTTTGTGTAGT 2700
QY 3030 AATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCA 3089
Db 2701 AATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCA 2760
QY 3090 GCATACCATTAAGAAATTTACTAAAAAATAA 3120
Db 2761 GCATACCATTAAGAAATTTACTAAAAAATAA 2791

RESULT 6
US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-2

Query Match      84.8%; Score 2644.4; DB 3; Length 2671;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 209 GCACATGCGCAGCAATGGAACAGAACAGACTGGGTGTGAGATATTTGAAACTCGG 268
Db 3 GTACCATGGCAGCAGCAATGGAACAGAACAGACTGGGTGTGAGATATTTGAAACTCGG 62
QY 269 ACTGTGAGGAGATATTTGAATCAGAGGATCGGCCCTAAATTTGAGCCTTTTATGTTGAGC 328
Db 63 ACTGTGAGGAGATATTTGAATCAGAGGATCGGCCCTAAATTTGAGCCTTTTATGTTGAGC 122
QY 329 GGTATTCCTGGAGTCAGCTTAAAGAGCTGTTCGCGATACAGAGAAATATCATGCTTACA 388
Db 123 GGTATTCCTGGAGTCAGCTTAAAGAGCTGTTCGCGATACAGAGAAATATCATGCTTACA 182
QY 389 TGATGGCTAAGCCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGACCTC 448
Db 183 TGATGGCTAAGCCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGACCTC 242
QY 449 ATTTCAGACAGAAATCTATTACCTTGGCATGCTGTGGTGAACAGAGAAAAATACACTGTTTT 508
Db 243 ATTTCAGACAGAAATCTATTACCTTGGCATGCTGTGGTGAACAGAGAAAAATACACTGTTTT 302
QY 509 ATTCTGAAATTCCTCAAAACTATCAATAGACGACGAGTCTTTAATGCTCTCTTTGGAAGCCTC 569
```

Db		303	ATTTGAAATTTCCCAAAACTATCAATAGACAGACTCTTAATGCTCTCTTGGAAAGCCTC	362	1383	GTTCCTCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCA	1442
Qy		569	TTTTGGATCTTTTTCAGGCAACACTGAGACTATGGAATGATTTCTCGAGAGAAAGAACTAT	628	1649	GTGGTGGCTCGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTA	1708
Db		363	TTTTGGATCTTTTTCAGGCAACACTGAGACTATGGAATGATTTCTCGAGAGAAAGAACTAT	422	1443	GTGGTGGCTCGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTA	1502
Qy		629	TAAGAGAAAGAAAACGCAATTTGGAAACAGTCGGAATTTGCTTTACGATATATCAACCAAGGAA	688	1709	CCAGTGGTGAATGGAGAGTTCTTGGCCGATCGATCTATATCCAAAGTTGATGAAGTCA	1768
Db		423	TAAGAGAAAGAAAACGCAATTTGGAAACAGTCGGAATTTGCTTTACGATATATCAACCAAGGAA	482	1503	CCAGTGGTGAATGGAGAGTTCTTGGCCGATCGATCTATATCCAAAGTTGATGAAGTCA	1562
Qy		689	GTGGAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTTATCACGTAAAGATGAGAGGCCAC	748	1769	GAAGCTGGTATATTTTGAAGGACCAAGAGCTCCCTCTTTAGAGCATCACCTGTAGCTAG	1828
Db		483	GTGGAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTTATCACGTAAAGATGAGAGGCCAC	542	1563	GAAGCTGGTATATTTTGAAGGACCAAGAGCTCCCTCTTTAGAGCATCACCTGTAGCTAG	1622
Qy		749	AAGGATTTTACGCAACCACTTTAAGGCCCAATCTAGTGGAAATAGTGTCTCCCAACATAC	808	1829	TCAGTTAGCTAAATCTCGAGAGGTGACAAAGGCTGACGACCGTGGCTACTACATCTCT	1888
Db		543	AAGGATTTTACGCAACCACTTTAAGGCCCAATCTAGTGGAAATAGTGTCTCCCAACATAC	602	1623	TCAGTTAGCTAAATCTCGAGAGGTGACAAAGGCTGACGACCGTGGCTACTACATCTCT	1682
Qy		809	GGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTATACATAGCAACG	868	1889	GCTGCATCAGTCAGCACCTGTGACTCTTTTATAAGTAAAGTATAGTAAACGAGAAATCCAC	1948
Db		603	GGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACG	662	1683	GCTGCATCAGTCAGCACCTGTGACTCTTTTATAAGTAAAGTATAGTAAACGAGAAATCCAC	1742
Qy		869	ATATTTGGATATCTAACATCGTRACAGAGAAAGAGAGACTCACTTATGTGCACAATG	928	1949	ACTGTGTCTCTTTTACAAAGCTATCAAGTCTCTGAAAGTGAACCCAACTTGCACAAACAAAG	2008
Db		663	ATATTTGGATATCTAACATCGTRACAGAGAAAGAGAGACTCACTTATGTGCACAATG	722	1743	ACTGTGTCTCTTTTACAAAGCTATCAAGTCTCTGAAAGTGAACCCAACTTGCACAAACAAAG	1802
Qy		929	AGCTAGCCAAATCGAGAGAGATCGAGATCAGCTGGAGTCCGTACCTTTGTTCTCCAAG	988	2009	AATTTTGGGCCACACATTTTGGATTCAGCAGGTCTCTCTCTGACTATCTCTCCAGCCTCATG	2068
Db		723	AGCTAGCCAAATCGAGAGAGATCGAGATCAGCTGGAGTCCGTACCTTTGTTCTCCAAG	782	1803	AATTTTGGGCCACACATTTTGGATTCAGCAGGTCTCTCTCTGACTATCTCTCCAGCCTCATG	1862
Qy		989	AAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACCTCCAGTG	1048	2069	TTTTCTCTTTTGAAGTACTACTCGATTTACATTTGATGGGATGCTCTCAAGACCTCATG	2128
Db		783	AAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACCTCCAGTG	842	1863	TTTTCTCTTTTGAAGTACTACTCGATTTACATTTGATGGGATGCTCTCAAGACCTCATG	1922
Qy		1049	GTGGTAAATCTTAGAATCTATATGAAGAAATAGATGAATCTGAGGTGGAAATTTATTC	1108	2129	ATCTACAGCTGGAAAGAAATATCTACTGTCTGTTTTCATATATGTTGGTCTCTCAGGTGC	2188
Db		843	GTGGTAAATCTTAGAATCTATATGAAGAAATAGATGAATCTGAGGTGGAAATTTATTC	902	1923	ATCTACAGCTGGAAAGAAATATCTACTGTCTGTTTTCATATATGTTGGTCTCTCAGGTGC	1982
Qy		1109	ATGTTACATCCCTATGTTGAAACAAAGGAGGCGAGATTCATTCGTTATCTCTAAACAG	1168	2189	AGTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGATACCTAGCCTCTC	2248
Db		903	ATGTTACATCCCTATGTTGAAACAAAGGAGGCGAGATTCATTCGTTATCTCTAAACAG	962	1983	AGTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGATACCTAGCCTCTC	2042
Qy		1169	GTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATATGATTTGATGCTGAAGGAA	1228	2249	TAGGTTATGTTGTTTGTAGTATAGCAACAGGGATCTCTGACCGAGGCTTAAATTTG	2308
Db		963	GTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATATGATTTGATGCTGAAGGAA	1022	2043	TAGGTTATGTTGTTTGTAGTATAGCAACAGGGATCTCTGACCGAGGCTTAAATTTG	2102
Qy		1229	GGATCATAGATGTCATAGATAGGAACTAAATTCACCTTTTGAGATTTCTATTGAGGAG	1288	2309	AAGGCGCTTTTAAATATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAAGGACTCC	2368
Db		1023	GGATCATAGATGTCATAGATAGGAACTAAATTCACCTTTTGAGATTTCTATTGAGGAG	1082	2103	AAGGCGCTTTTAAATATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAAGGACTCC	2162
Qy		1289	TTGAATATTTGCCAGAGCTGGATGGACTCTGAGGGAAATATGCTTGGTCCATCCTAC	1142	2369	AATATCTAGCTTCTCGATATGATTTCAATTTGACTTTAGATCGTGTGGGCATCCAGGCTGGT	2428
Db		1083	TTGAATATTTGCCAGAGCTGGATGGACTCTGAGGGAAATATGCTTGGTCCATCCTAC	1142	2163	AATATCTAGCTTCTCGATATGATTTCAATTTGACTTTAGATCGTGTGGGCATCCAGGCTGGT	2222
Qy		1349	TAGATCGCTCCGAGCTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTCCAG	1408	2429	CCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGATATCTTCAGGGTTG	2488
Db		1143	TAGATCGCTCCGAGCTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTCCAG	1202	2223	CCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGATATCTTCAGGGTTG	2282
Qy		1409	TAGAAATGATGTTTATGAAAGGAGAGACTGATGAGTCAGTCCGCTGATCTCTGACGC	1468	2489	CTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTT	2548
Db		1203	TAGAAATGATGTTTATGAAAGGAGAGACTGATGAGTCAGTCCGCTGATCTCTGACGC	1262	2283	CTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTT	2342
Qy		1469	CACTAATTTATCTAGAGAAACACAGACATCTGGATTAATATCCATGATCTTTTCATG	1528	2549	ATATGGGTCACTCCACAGAAATGAACAGGGCTAATTTAGGATCTGTGGCCATGCAAG	2608
Db		1263	CACTAATTTATCTAGAGAAACACAGACATCTGGATTAATATCCATGATCTTTTCATG	1322	2343	ATATGGGTCACTCCACAGAAATGAACAGGGCTAATTTAGGATCTGTGGCCATGCAAG	2402
Qy		1529	TTTTTCCCAAGTCAAGAGGAAATTTGATTTTATTTTGGCTCTCTGAAATGCAAAACAG	1588	2609	CAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGATGAGA	2668
Db		1323	TTTTTCCCAAGTCAAGAGGAAATTTGATTTTATTTTGGCTCTCTGAAATGCAAAACAG	1382	2403	CAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGATGAGA	2462
Qy		1589	GTTCCTCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCA	1648	2669	ATGTCATTTTGAACATAGCAGTATATCTAGTGTGTTTAAAGTGGGCTGGAAGGCAAT	2728
					2463	ATGTCATTTTGCACATACAGTATATTTACTAGTGTGTTTAAAGTGGGCTGGAAGGCAAT	2522

Qy	2729	ATGATTACAGATCTATCTCTCAGAGAGAGACACAGCATTAAGAGTTTCTGTAATCGGGAGAAC	2789
Db	2523	ATGATTTACAGATCTATCTCTCAGAGAGAGACACAGCATTAAGAGTTCTGTAATCGGGAGAAC	2582
Qy	2789	ATTATCAACTGCATCTTTTGCACACTACCTTCAAGAAAAACCTTGGATCAGGTATTGCTGCTC	2848
Db	2583	ATTATGAACATGCATCTTTTGCACACTACCTTCAAGAAAAACCTTGGATCAGGTATTGCTGCTC	2642
Qy	2849	TAAAGTGTATATAA	2862
Db	2643	TAAAGTGTATAGA	2656

RESULT 7
 US-09-976-674-8
 ; Sequence 8, Application US/09976674
 ; Patent No. 6844180
 ; GENERAL INFORMATION:
 ; APPLICANT: Qi, Steve
 ; APPLICANT: Akinsanya, Karen
 ; APPLICANT: Riviere, Pierre
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
 ; FILE REFERENCE: 70669
 ; CURRENT APPLICATION NUMBER: US/09/976,674
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: US 60/240,117
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 4523
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-976-674-8

Qy	481	GGTGAGAA	CAGAGAAAAT	TACACTGTTT	TATCTGAAAT	TCCAAAAC	TATCAATAGACCA	540
Db	481	GGTGAGAA	CAGAGAAAAT	TACACTGTTT	TATCTGAAAT	TCCAAAAC	TATCAATAGACCA	540
Qy	541	CGAGCTTAA	TGCTCTCTTGG	AGCCCTCTTTTGG	ATCTTTTTCAGG	CAACACTCGACTAT	600	
Db	541	CGAGCTTAA	TGCTCTCTTGG	AGCCCTCTTTTGG	ATCTTTTTCAGG	CAACACTCGACTAT	600	
Qy	601	GGAAATGTA	TCTCTCGAGA	AGAGAACTATTA	AGAGAAAGAAA	CGCATTTGGA	CAAGTCGGA	660
Db	601	GGAAATGTA	TCTCTCGAGA	AGAGAACTATTA	AGAGAAAGAAA	CGCATTTGGA	CAAGTCGGA	660
Qy	661	ATTGCTCTT	TACGATTATC	CAACAGNAGT	GGNACATTTCTG	TTTCAAGC	CGGTAGTGG	720
Db	661	ATTGCTCTT	TACGATTATC	CAACAGNAGT	GGNACATTTCTG	TTTCAAGC	CGGTAGTGG	720
Qy	721	ATTTATCAG	GTAAAGATG	GAGGGCCAC	CAAGGATTTTAC	GCAACAACTTTA	AGGCCCAAT	780
Db	721	ATTTATCAG	GTAAAGATG	GAGGGCCAC	CAAGGATTTTAC	GCAACAACTTTA	AGGCCCAAT	780
Qy	781	CTAGTGGAA	CTAGTTGT	TCCAA	CATACGGATGGAT	TCCAAAATTA	TATGCCCCGTGATCCA	840
Db	781	CTAGTGGAA	CTAGTTGT	TCCAA	CATACGGATGGAT	TCCAAAATTA	TATGCCCCGTGATCCA	840
Qy	841	GACTGGAT	TGCTTTTAT	CATAGCA	CGATATTTGG	GATCTTAA	CATGTTACCAAGAA	900
Db	841	GACTGGAT	TGCTTTTAT	CATAGCA	CGATATTTGG	GATCTTAA	CATGTTACCAAGAA	900
Qy	901	GAAGGAGAC	TCACTTAT	TGTGCA	CAATGAGCTAG	CCAA	CATGGAAGAGATGCCAGATCA	960
Db	901	GAAGGAGAC	TCACTTAT	TGTGCA	CAATGAGCTAG	CCAA	CATGGAAGAGATGCCAGATCA	960
Qy	961	GCTGGAGT	CGCTAC	CTTTGTTCT	CCAAAGAAATTTG	ATAGATATTTCT	GGCTATTTGGTGG	1020
Db	961	GCTGGAGT	CGCTAC	CTTTGTTCT	CCAAAGAAATTTG	ATAGATATTTCT	GGCTATTTGGTGG	1020
Qy	1021	TGTCAAAA	AGCTGAA	AAACAACTCC	CAAGTGGTG	TATAAATTTCT	TAGAAATTTCTATACAAAGAA	1080
Db	1021	TGTCAAAA	AGCTGAA	AAACAACTCC	CAAGTGGTG	TATAAATTTCT	TAGAAATTTCTATACAAAGAA	1080
Qy	1081	AATGATGAT	CTGAGGT	GGAATTTAT	CATGTTACAT	TCCCTATGTT	GAAACAAAGGAGG	1140
Db	1081	AATGATGAT	CTGAGGT	GGAATTTAT	CATGTTACAT	TCCCTATGTT	GAAACAAAGGAGG	1140
Qy	1141	GCAGATTCA	TTCGGTTAT	TCTTAA	ACAGGTACAGCAAA	TCTTAA	AGTCACATTTTAAAGATG	1200
Db	1141	GCAGATTCA	TTCGGTTAT	TCTTAA	ACAGGTACAGCAAA	TCTTAA	AGTCACATTTTAAAGATG	1200
Qy	1201	TCAGAAAT	AATGATG	CTGTGA	AGGAAGGATCAT	GATGTG	CATAGTAAGGAACATAAT	1260
Db	1201	TCAGAAAT	AATGATG	CTGTGA	AGGAAGGATCAT	GATGTG	CATAGTAAGGAACATAAT	1260
Qy	1261	CAACCTTT	TGAGATTT	CTAATTTGA	AGGATTTGA	ATATATTTG	CCAGAGCTGGATGACTCCT	1320
Db	1261	CAACCTTT	TGAGATTT	CTAATTTGA	AGGATTTGA	ATATATTTG	CCAGAGCTGGATGACTCCT	1320
Qy	1321	GAGGAAAA	TATGCTT	TGCTCAT	CTACTAGATCG	CTCCAGACTCG	CTACAGATAGTG	1380
Db	1321	GAGGAAAA	TATGCTT	TGCTCAT	CTACTAGATCG	CTCCAGACTCG	CTACAGATAGTG	1380
Qy	1381	TTGATCTCA	CTGAAATTTAT	TCCAGT	TAGAAGATG	TGTTATG	AAAGGCAGAGATCT	1440
Db	1381	TTGATCTCA	CTGAAATTTAT	TCCAGT	TAGAAGATG	TGTTATG	AAAGGCAGAGATCT	1440
Qy	1441	ATTGAGT	CTAGTGG	TATCTGT	GACGACATAAT	TATCTATG	AAACAAACACAGATC	1500
Db	1441	ATTGAGT	CTAGTGG	TATCTGT	GACGACATAAT	TATCTATG	AAACAAACACAGATC	1500
Qy	1501	TGGATAAA	TATCAT	GACATCTTT	CATGTTTTT	CCCCCAAGTCA	CGAAGAGGAAATTTGAG	1560
Db	1501	TGGATAAA	TATCAT	GACATCTTT	CATGTTTTT	CCCCCAAGTCA	CGAAGAGGAAATTTGAG	1560
Qy	1561	TTTATTTT	TTTGGCTCT	GAAATG	CAAAACAGGTTT	CGGTCA	TTTATACAAATATCATCTATT	1620

Db 1561 TTTATTTTGTCTCGAATGCAAAACAGGTTTCCTGTCATTTATACAAAATACATCATTT 1620
Qy 1621 TTAAGGAAGCAAAATATAACGATCCAGTGGTGGCTGCTCCAGTGAATTTCAAG 1680
Db 1621 TTAAGGAAGCAAAATATAACGATCCAGTGGTGGCTGCTCCAGTGAATTTCAAG 1680
Qy 1681 TGTCTATCAAAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db 1681 TGTCTATCAAAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Qy 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTCGTATATTTTGAAGGCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTCGTATATTTTGAAGGCACCAAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACCTGTAGTACGTAGTACGTAAATCCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTAGTACGTAGTACGTAAATCCTGGAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACCAAGAGATCCACACTGTGTCTGCTTTTCAAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAGATCCACACTGTGTCTGCTTTTCAAGCTATCAAGTCT 1980
Qy 1981 GAAGATGACCAACTGTCAGAAACAAAGGAATTTTGGGCCACATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCAACTGTCAGAAACAAAGGAATTTTGGGCCACATTTTGGATTCAGCAGGT 2040
Qy 2041 CTTCTTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTTACA 2100
Db 2041 CTTCTTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTTACA 2100
Qy 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG 2160
Qy 2161 CTGTTTCATATATGGTGGTCTCAGTGCAGTGTGTGAATAATCGGTTTAAAGGATCAAG 2220
Db 2161 CTGTTTCATATATGGTGGTCTCAGTGCAGTGTGTGAATAATCGGTTTAAAGGATCAAG 2220
Qy 2221 TATTTCCGCTTGAAATACCTAGCTCTAGGTTATGTGTTTGTAGTATAGACACAGG 2280
Db 2180 ----- 2179
Qy 2281 GGATCCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAATGGGTCAAATA 2340
Db 2180 ----- 2179
Qy 2341 GAAATTGAGATCAGGTGGAAGGACTCAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
Db 2180 ----- 2179
Qy 2401 TTAGATCGGTGGGCATCCAGCGTGGTGCCTATGGAGGATACCTCTCCCTGATGGCATTA 2460
Db 2180 ----- 2179
Qy 2461 ATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATC 2520
Db 2180 -----GGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATC 2216
Qy 2521 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACCAAGATGAACAGGGC 2580
Db 2217 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACCAAGATGAACAGGGC 2276
Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTA 2640
Db 2277 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTA 2336
Qy 2641 CTGCTCTTACATGTTTCTCTGATGAGATGTCATTTTGCATACCAATGATTTACTG 2700

Db 2337 CTGCTCTTACATGGTTTCTCGGATGAGAAATGTCATTTTGCACATACCATATATCTG 2396
Qy 2701 AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCCTCAGGAGACAC 2760
Db 2397 AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCCTCAGGAGACAC 2456
Qy 2761 AGCATAGAGTTCCTGAAATCGGAGAACATATGAACTGCACTTTTGCACCTTCAA 2820
Db 2457 AGCATAGAGTTCCTGAAATCGGAGAACATATGAACTGCACTTTTGCACCTTCAA 2516
Qy 2821 GAAACCTTGGATCACGCTATTGCTCTCTAAAGTGAATATAATTTTGAACCTGTTGAGAAC 2880
Db 2517 GAAACCTTGGATCACGCTATTGCTCTCTAAAGTGAATATAATTTTGAACCTGTTGAGAAC 2576
Qy 2881 TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGTTTAAATCAACAGAAACACAGA 2940
Db 2577 TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGTTTAAATCAACAGAAACACAGA 2636
Qy 2941 ATTGATCATACATTTTGTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
Db 2637 ATTGATCATACATTTTGTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2696
Qy 3001 CCATGCGGGTCTACGGTTTGTGTAGTAAATCTAACTTAAACCCCATCATCTCAAAA 3060
Db 2697 CCATGCGGGTCTACGGTTTGTGTAGTAAATCTAACTTAAACCCCATCATCTCAAAA 2756
Qy 3061 TCAAAATGATACATATCTCTGAGAGACCCCAATACCATAAGAAATTAATAAAAAA 3120
Db 2757 TCAAAATGATACATATCTCTGAGAGACCCCAATACCATAAGAAATTAATAAAAAA 2816

RESULT 8
US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976.674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-14

Query Match 65.9%; Score 2056.4; DB 3; Length 4309;
Best Local Similarity 83.2%; Pred. No. 0;
Matches 2599; Conservative 0; Mismatches 1; Indels 522; Gaps 2;
Qy 1 AAGTGTAAAGCTTCGAGGCGAAGCGCGCTGCTACTGCGCGCGCTCTTCTTAGTGCCG 60
Db 1 AAGTGTAAAGCTTCGAGGCGAAGCGCGCTGCTACTGCGCGCGCTCTTCTTAGTGCCG 60
Qy 61 CGTTCGCCCTGGGTGTGTCACCGCGCGCGCGCGAGAGCCACTGCAACCGAGCCG 120
Db 61 CGTTCGCCCTGGGTGTGTCACCGCGCGCGCGCGAGAGCCACTGCAACCGAGCCG 120
Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGTTCGGGACGG 180
Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGTTCGGGACGG 180
Qy 181 TCCGGCGCGCGCGCGGAGAAAATGCAACATGGCAGCAGCAATGGAAACAGACAG 240
Db 181 TCCGGCGCGCGCGCGGAGAAAATGCAACATGGCAGCAGCAATGGAAACAGACAG 240

Qy	241	CTGGGTGTTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGG	300
Ds	241	CTGGGTGTTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGG	300
Qy	301	CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCTTGGAGTCAGCTTAAAGAGCTGCTT	360
Ds	301	CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCTTGGAGTCAGCTTAAAGAGCTGCTT	360
Qy	361	GCCGATPACCAAGAAATATCATGGCTACATGATGGCTAAGGCCACACATGATTTTCATGTTT	420
Ds	361	GCCGATPACCAAGAAATATCATGGCTACATGATGGCTAAGGCCACACATGATTTTCATGTTT	420
Qy	421	GTGAAGAGAAATGATCCAGATGAGACCTCATTTACAGACAGAAATCTATTACCTTGGCATGTCT	480
Ds	421	GTGAAGAGAAATGATCCAGATGAGACCTCATTTACAGACAGAAATCTATTACCTTGGCATGTCT	480
Qy	481	GGTGAGAACAGAGAAATACACTGTTTTTATCTGAATTCCTCAAAATCCCAAACTATCAATAGACGA	540
Ds	481	GGTGAGAACAGAGAAATACACTGTTTTTATCTGAATTCCTCAAAATCCCAAACTATCAATAGACGA	540
Qy	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Ds	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGAAATGTAATTCGAGAGAGAACTATTAAAGAGAGAAAGAAACGCAATTTGGAACAGTCGGA	660
Ds	601	GGAAATGTAATTCGAGAGAGAACTATTAAAGAGAGAAAGAAACGCAATTTGGAACAGTCGGA	660
Qy	661	ATTGCTTCTTACGATTTATCACCAAGGAAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA	720
Ds	661	ATTGCTTCTTACGATTTATCACCAAGGAAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA	720
Qy	721	ATTATACAGTAAAGATGAGGGCCACAAAGATTTAGCGAAACAACTTTAAGGCCCAAT	780
Ds	721	ATTATACAGTAAAGATGAGGGCCACAAAGATTTAGCGAAACAACTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGATCCA	840
Ds	781	CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Ds	841	GACTGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Qy	901	GAAAGAGACTCACTTATGTGCAATGAGCTAGCCACATGGAAGAGATGCCAGATCA	960
Ds	901	GAAAGAGAGACTCACTTATGTGCAATGAGCTAGCCACATGGAAGAGATGCCAGATCA	960
Qy	961	GCTGGAGTGGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTGGCTATTGGTGG	1020
Ds	961	GCTGGAGTGGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTGGCTATTGGTGG	1020
Qy	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTGAATTTCTATATGAGAA	1080
Ds	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTGAATTTCTATATGAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGAAATTTATCATGTTACATCCCTATGTTGGAAACAGGAGG	1140
Ds	1081	AATGATGAATCTGAGGTGAAATTTATCATGTTACATCCCTATGTTGGAAACAGGAGG	1140
Qy	1141	GCAGATTCATTCGGTTATCCTAAACAGGTACAGCAATCTTAAGTCACCTTTTAAGATG	1200
Ds	1141	GCAGATTCATTCGGTTATCCTAAACAGGTACAGCAATCTTAAGTCACCTTTTAAGATG	1200
Qy	1201	TCAGAAATATGATTCATGCTGAAGAGATCATATGATGTCATAGATATAGGAACATAAT	1260
Ds	1201	TCAGAAATATGATTCATGCTGAAGAGATCATATGATGTCATAGATATAGGAACATAAT	1260
Qy	1261	CAACCTTTTGAGATTCATTTGAAGAGTGTGAATATATGCGAGAGCTGGATGACTCTCT	1320
Ds	1261	CAACCTTTTGAGATTCATTTGAAGAGTGTGAATATATGCGAGAGCTGGATGACTCTCT	1320
Qy	1230	-----	1230
Ds	1230	-----	1230
Qy	1321	GAGGAAAAATATGCTTGGTCCATCCTACTAGATCGTCCCAGACTCGCCTACAGATAGTG	1380
Ds	1321	GAGGAAAAATATGCTTGGTCCATCCTACTAGATCGTCCCAGACTCGCCTACAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Ds	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Qy	1441	ATTGAGTCAGTCGCTGATTTCTGTGACGCCACTAATTATCTATATGAAGAAACAAAGACATC	1500
Ds	1441	ATTGAGTCAGTCGCTGATTTCTGTGACGCCACTAATTATCTATATGAAGAAACAAAGACATC	1500
Qy	1501	TGATATAATCCATGACATCTTTTCATGTTTTTCCCCAAAGTCACAGAGGAAATTTGAG	1560
Ds	1501	TGATATAATCCATGACATCTTTTCATGTTTTTCCCCAAAGTCACAGAGGAAATTTGAG	1560
Qy	1561	TTTTATTTTCTCTGAAATGCAAAACAGGTTTTCGTCATTTTATACAAAATTTACATCTATT	1620
Ds	1561	TTTTATTTTCTCTGAAATGCAAAACAGGTTTTCGTCATTTTATACAAAATTTACATCTATT	1620
Qy	1621	TTAAAGGAAAGCAAAATATAAAGCATCCAGTGGTGGGCTGCCTGCCAAGTGATTTCAAG	1680
Ds	1621	TTAAAGGAAAGCAAAATATAAAGCATCCAGTGGTGGGCTGCCTGCCAAGTGATTTCAAG	1680
Qy	1681	TGTCCTTCAAGAGAGAGATAGCAATTACCAGTGGTGAATGGGAAGTTCTTGGCCGCGAT	1740
Ds	1681	TGTCCTTCAAGAGAGAGATAGCAATTACCAGTGGTGAATGGGAAGTTCTTGGCCGCGAT	1740
Qy	1741	GGATCTAATATCCAGTTGATGAGTCCAGAGGCTGGTATATTTTGAAGGCCACCAAGAGAC	1800
Ds	1741	GGATCTAATATCCAGTTGATGAGTCCAGAGGCTGGTATATTTTGAAGGCCACCAAGAGAC	1800
Qy	1801	TCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAGTAAATCTCTGGAGAGGTGCAAGG	1860
Ds	1801	TCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAGTAAATCTCTGGAGAGGTGCAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTTCTGTCATCAGTCAGCACTGTGACTCTTTTATA	1920
Ds	1861	CTGACTGACCGTGGCTACTCACATTTCTTCTGTCATCAGTCAGCACTGTGACTCTTTTATA	1920
Qy	1921	AGTAAGTATAGTAACAGAGAAATCCACATGTTGTCCTCTTTTACAGCTATCAAGTCCT	1980
Ds	1921	AGTAAGTATAGTAACAGAGAAATCCACATGTTGTCCTCTTTTACAGCTATCAAGTCCT	1980
Qy	1981	GAAAGTGAACCAACTTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
Ds	1981	GAAAGTGAACCAACTTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTTACA	2100
Ds	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTTACA	2100
Qy	2101	TTGTATGGAGTGTCTTACAGGCTCATGATCTACAGCTGGAAGAAATATCTTACTGTG	2160
Ds	2101	TTGTATGGAGTGTCTTACAGGCTCATGATCTACAGCTGGAAGAAATATCTTACTGTG	2160
Qy	2161	CTGTTTCATATATGGTGG--TCCTCAGGTGCGAGTGGTGAATAATCGGTTTAAAGAGTCA	2218
Ds	2161	CTGTTTCATATATGGTGG--TCCTCAGGTGCGAGTGGTGAATAATCGGTTTAAAGAGTCA	2218
Qy	2219	AGTATTTCCGCTTGAATACCCTAGCCTCTCTAGGTTATGTGGTGTAGTATAGACAACA	2278
Ds	2219	AGTATTTCCGCTTGAATACCCTAGCCTCTCTAGGTTATGTGGTGTAGTATAGACAACA	2278
Qy	2279	GGGATCCTGTCAACGAGGCTTTAAATTTGAAGGGCCCTTTAAATATATAAATGGGTCAA	2338
Ds	2279	GGGATCCTGTCAACGAGGCTTTAAATTTGAAGGGCCCTTTAAATATATAAATGGGTCAA	2338
Qy	2339	TAGAAATTTGAGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTG	2398
Ds	2339	TAGAAATTTGAGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTG	2398
Qy	2399	ACTTAGATCGTGTGGGCAATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAT	2458
Ds	2399	ACTTAGATCGTGTGGGCAATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAT	2458

Db 1881 ACTTAGATCGTGTGGGATCCAGGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCAT 1940
QY 2459 TAATGCGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGA 2518
Db 1941 TAATGCGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGA 2000
QY 2519 TCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCACTGACAGGAATGAACAGG 2578
Db 2001 TCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCACTGACAGGAATGAACAGG 2060
QY 2579 GCTATTACTTAGGATCTGTGGCATCAAGCAAAAGTTCCCTCTGAAACCAATCGTT 2638
Db 2061 GCTATTACTTAGGATCTGTGGCATCAAGCAAAAGTTCCCTCTGAAACCAATCGTT 2120
QY 2639 TACTGCTCTTACATGCTTCTGTGATGAGAACTGTCATTTGCAATACCAAGTATATTAC 2698
Db 2121 TACTGCTCTTACATGCTTCTGTGATGAGAACTGTCATTTGCAATACCAAGTATATTAC 2180
QY 2699 TGAGTTTTTATGAGGCTGGAAAGCCATATATGATTTACAGATCTATCTCAGAGAGAC 2758
Db 2181 TGAGTTTTTATGAGGCTGGAAAGCCATATATGATTTACAGATCTATCTCAGAGAGAC 2240
QY 2759 ACAGATAAGAGTTCTGAAATCGGAGAACATTTAGAACTGCACTCTTTGCACTACCTTC 2818
Db 2241 ACAGATAAGAGTTCTGAAATCGGAGAACATTTAGAACTGCACTCTTTGCACTACCTTC 2300
QY 2819 AAGAAACCTTTGGATCAGTATCTGCTCTAAAGAGTATATATTTGACCTGTGTAGA 2878
Db 2301 AAGAAACCTTTGGATCAGTATCTGCTCTAAAGAGTATATATTTGACCTGTGTAGA 2360
QY 2879 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACA 2938
Db 2361 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACA 2420
QY 2939 GAATTTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAATTAATGTGG 2998
Db 2421 GAATTTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAATTAATGTGG 2480
QY 2999 TGCCATGAGGGCTCAGGTTTGTGTAGTATCTAATCTAATACCTTAACCCCACTGCTCAA 3058
Db 2481 TGCCATGAGGGCTCAGGTTTGTGTAGTATCTAATCTAATACCTTAACCCCACTGCTCAA 2540
QY 3059 ATCAATGATACATATCTCGAGAGACCCAGCAATACCAAGAAATTAATAAAAAA 3118
Db 2541 ATCAATGATACATATCTCGAGAGACCCAGCAATACCAAGAAATTAATAAAAAA 2600
QY 3119 AA 3120
Db 2601 AA 2602

RESULT 9
US-10-070-464-6
; Sequence 6, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070.464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1669

; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-6
Query Match 43.2%; Score 1347; DB 3; Length 1669;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 0; Indels 300; Gaps 1;
QY 1164 AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAGATGTCAGAAATTAATGATTGATGCTGA 1223
Db 1 AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAGATGTCAGAAATTAATGATTGATGCTGA 60
QY 1224 AGGAAGGATCATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAATTCATATTGA 1283
Db 61 AGGAAGGATCATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAATTCATATTGA 120
QY 1284 AGGAGTTCGAAATATATGTCAGAGCTGGATGGATCTCTGAGGGGAAAAATATGCTTGGTCCAT 1343
Db 121 AGGAGTTCGAAATATATGTCAGAGCTGGATGGATCTCTGAGGGGAAAAATATGCTTGGTCCAT 180
QY 1344 CCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGTGATCTCACCTGAATTAATTTAT 1403
Db 181 CCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGTGATCTCACCTGAATTAATTTAT 240
QY 1404 CCCAGTACAGATGATGTTATGGAAGCGCAGAGACTCATTGAGTCACTGCTGATTCCTG 1463
Db 241 CCCAGTACAGATGATGTTATGGAAGCGCAGAGACTCATTGAGTCACTGCTGATTCCTG 300
QY 1464 GACGCACTAATTAATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1523
Db 301 GACGCACTAATTAATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
QY 1524 TCATGTTTTTCCCAAGTACGAAAGGAAATGAGTTTATTTTGGCTCTGAAATGCAA 1583
Db 361 TCATGTTTTTCCCAAGTACGAAAGGAAATGAGTTTATTTTGGCTCTGAAATGCAA 420
QY 1584 AACAGGTTTCGGTCAATTTATACAAATTAATCATCTATTTTAAAGAAAGCAATATATAACG 1643
Db 421 AACAGGTTTCGGTCAATTTATACAAATTAATCATCTATTTTAAAGAAAGCAATATATAACG 480
QY 1644 ATCCAGTGTGGGCTGCTGCTCCAAAGTGAITTCAGTGTCTTATCAAGAGAGATAGC 1703
Db 481 ATCCAGTGTGGGCTGCTGCTCCAAAGTGAITTCAGTGTCTTATCAAGAGAGATAGC 540
QY 1704 AATTACAGTGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATAATCCAAGTTGATGA 1763
Db 541 AATTACAGTGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATAATCCAAGTTGATGA 600
QY 1764 AGTCAGAAGGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTTTAGAGCATCACCTGTA 1823
Db 601 AGTCAGAAGGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTTTAGAGCATCACCTGTA 660
QY 1824 CGTAGTCACTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACCGTGGCTACTCACA 1883
Db 661 CGTAGTCACTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACCGTGGCTACTCACA 720
QY 1884 TTCTTCTGCATCAGTCACTGAGCTGAGCTTCTTTTATAGTAAGTATAGTAACCAAGAA 1943
Db 721 TTCTTCTGCATCAGTCACTGAGCTGAGCTTCTTTTATAGTAAGTATAGTAACCAAGAA 780
QY 1944 TCACACATGTGTGCTCTTTTACAGCTATCAAGTCTGAGTCCCTGAAAGTCAACCACTTGCRAAAC 2003
Db 781 TCACACATGTGTGCTCTTTTACAGCTATCAAGTCTGAGTCCCTGAAAGTCAACCACTTGCRAAAC 840
QY 2004 AAAGGAATTTTGGGCCACCATTTTGGATTGAGAGGCTCTTCTCTGATCTACTCTCC 2063
Db 841 AAAGGAATTTTGGGCCACCATTTTGGATTGAGAGGCTCTTCTCTGATCTACTCTCC 900
QY 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTTATGGGATGCTCTCAAGCC 2123
Db 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTTATGGGATGCTCTCAAGCC 960
QY 2124 TCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTCATATATATGTTGCTCTCA 2183

|||||
Db 961 TCATGATCTACAGCTGGAAGAAATATCTTACTGTCTGTTTCATATATATGGTGGCTCTCA 1020
QY 2184 GGTGCAAGTTGGTGAATAATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGC 2243
Db 1021 ----- 1020
QY 2244 CTCTCTAGTTATGTTGTTGATGATAGACAAACAGGGATCTGTCTACCGAGGGCTTAA 2303
Db 1021 ----- 1020
QY 2304 ATTTGAAGCGCTTTTAAATATATAAATGGGTCAAATAGAAATTCACGATCAGGTGGAAG 2363
Db 1021 ----- 1020
QY 2364 ACTCCAATATCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGG 2423
Db 1021 ----- 1020
QY 2424 CTGGTCTATGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGTCAAGATCTTTCAG 2483
Db 1021 ----- 1020
QY 2484 GGTGCTATGCTGGGGCCCCAGTCACTCTGTGTGATCTTCTATGATACAGGATACACGGA 2543
Db 1021 GGTGCTATGCTGGGGCCCCAGTCACTCTGTGTGATCTTCTATGATACAGGATACACGGA 1080
QY 2544 AGTTATATGGGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2603
Db 1081 AGTTATATGGGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
QY 2604 GCAAGCAGAAAAGTTCCCTCTGTAACCAATCGTTTACTGCTCTTACATGTTTCTCTGA 2663
Db 1141 GCAAGCAGAAAAGTTCCCTCTGTAACCAATCGTTTACTGCTCTTACATGTTTCTCTGA 1200
QY 2664 TGAGATGTCCTATTTTGACATACCAAGTATTAATCTGAGTTTCTTGTAGTGGGCTGAAA 2723
Db 1201 TGAGATGTCCTATTTTGACATACCAAGTATTAATCTGAGTTTCTTGTAGTGGGCTGAAA 1260
QY 2724 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAAGATTCCTGAAATCGGG 2783
Db 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAAGATTCCTGAAATCGGG 1320
QY 2784 AGAACAATTATGAATGCACTTTTGGCACTCTTTCACATCTCTTCAAGAAAACCTTGGATCAGTATTGC 2843
Db 1321 AGAACAATTATGAATGCACTTTTGGCACTCTTTCACATCTCTTCAAGAAAACCTTGGATCAGTATTGC 1380
QY 2844 TGCTCTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGATATACACTGGCTATT 2903
Db 1381 TGCTCTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGATATACACTGGCTATT 1440
QY 2904 AACCAAAATGAGAGGTTTAAATCAACAGAAAACACAGAAATTCATCATCATCATTTTGTATCC 2963
Db 1441 AACCAAAATGAGAGGTTTAAATCAACAGAAAACACAGAAATTCATCATCATCATTTTGTATCC 1500
QY 2964 TGCCATGTAACATCTACTCTGAAAATAAATGTGGTCCATGACAGGGGTCTACGGTTTGT 3023
Db 1501 TGCCATGTAACATCTACTCTGAAAATAAATGTGGTCCATGACAGGGGTCTACGGTTTGT 1560
QY 3024 GGTAGTAATCTAATACCTTAACCCCAATGCTTCAAAATCAAAATGATACATATTTCTTGAGA 3083
Db 1561 GGTAGTAATCTAATACCTTAACCCCAATGCTTCAAAATCAAAATGATACATATTTCTTGAGA 1620
QY 3084 GACCCAGCAATACCAAGAATTTACTAAAAA 3120
Db 1621 GACCCAGCAATACCAAGAATTTACTAAAAA 1657

RESULT 10

US-10-070-464-4

; Sequence 4, Application US/10070464

; Patent No. 6881564

; GENERAL INFORMATION:

; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-4

Query Match 33.2%; Score 1036; DB 3; Length 1197;

Best Local Similarity 89.0%; Pred. No. 1.6e-293;

Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1781 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCCTGCTAGTACGTTACGTAA 1840
Db 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCCTGCTAGTACGTTACGTAA 60
QY 1841 ATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGCATCAGTC 1900
Db 61 ATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGCATCAGTC 120
QY 1901 AGCACTGTGACTTCTTTATAAGTATAGTAACACAGAGAAATCCACACTGTGTGCC 1960
Db 121 AGCACTGTGACTTCTTTATAAGTATAGTAACACAGAGAAATCCACACTGTGTGCC 180
QY 1961 TTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGGAAATTTTGGGCCA 2020
Db 181 TTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGGAAATTTTGGGCCA 240
QY 2021 CCATTTTGGATTACAGAGTCTCTTCTGTGATATATCTTCCAGAAATTTTCTCTTTG 2080
Db 241 CCATTTTGGATTACAGAGTCTCTTCTGTGATATATCTTCCAGAAATTTTCTCTTTG 300
QY 2081 AAAGTACTACTGATTTTCAITTTGATGGATGCTCTCAAGCTCATGATCTACAGCTG 2140
Db 301 AAAGTACTACTGATTTTCAITTTGATGGATGCTCTCAAGCTCATGATCTACAGCTG 360
QY 2141 GAAAGAAATATCTTACTGTCTGTTTCATATATGGTGGTCTCCTCAGTGCAGTTGGTGAATA 2200
Db 361 GAAAGAAATATCTTACTGTCTGTTTCATATATGGTGGTCTCCTCAGTGCAGTTGGTGGTCCA 403
QY 2201 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGTTATGTGG 2260
Db 404 ----- 403
QY 2261 TTGTAGTATAGACAAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGCGGCTTTTA 2320
Db 404 ----- 403
QY 2321 AATATAAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAAAGGACTCCAATATCTAGCTT 2380
Db 404 -----GGTCAAAATAGAAATTTGACGATCAGGTGGAAAGGACTCCAATATCTAGCTT 453
QY 2381 CTCGATATGATTTTCAITTTGATCGTGTGGGCAATCCACGGTGTGCTTATGAGGAT 2440
Db 454 CTCGATATGATTTTCAITTTGATCGTGTGGGCAATCCACGGTGTGCTTATGAGGAT 513
QY 2441 ACCTTCCCTGATGGCAATTAATGCAGAGTCAAGATATCTTCAAGGTTGCTATTCCTGGG 2500
Db 514 ACCTTCCCTGATGGCAATTAATGCAGAGTCAAGATATCTTCAAGGTTGCTATTCCTGGG 573
QY 2501 CCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACC 2560

Db	574	CCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAAAGTTATATGGTCAAC	633
Qy	2561	CTGACGAGATGAACAGGGCTATCTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC	2620
Db	634	CTGACGAGATGAACAGGGCTATCTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC	693
Qy	2621	CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTGTGATGAGAAATGTCATTTTG	2680
Db	694	CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTGTGATGAGAAATGTCATTTTG	753
Qy	2691	CACATACCAAGTATATTAAGTGTCTTGTAGTGGGCTGAAAGCCATATGATTTACAGA	2740
Db	754	CACATACCAAGTATATTAAGTGTCTTGTAGTGGGCTGAAAGCCATATGATTTACAGA	813
Qy	2741	TCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGGGAAACATTTGAACCTGC	2800
Db	814	TCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGGGAAACATTTGAACCTGC	873
Qy	2801	ATCTTTTGCACCTCTCAAGAAAACCTTGGATCAGTATTTGCTCTCTAAAAGTGATAT	2860
Db	874	ATCTTTTGCACCTCTCAAGAAAACCTTGGATCAGTATTTGCTCTCTAAAAGTGATAT	933
Qy	2861	AAATTTGACCTGTAGAACTCTCTGTATACACTGGCTATTTAACCAAAATGAGGAGTT	2920
Db	934	AAATTTGACCTGTAGAACTCTCTGTATACACTGGCTATTTAACCAAAATGAGGAGTT	993
Qy	2921	TAATCAACAGAAAACAGAAATTTGATCAGTATTTGATCCTGCCATGTAAACATCTAC	2980
Db	994	TAATCAACAGAAAACAGAAATTTGATCAGTATTTGATCCTGCCATGTAAACATCTAC	1053
Qy	2981	TCCTGAAAATAAATGTGGTGCATGACGGGTCTACGGTTTGTGGTAGTAATCTTAATACC	3040
Db	1054	TCCTGAAAATAAATGTGGTGCATGACGGGTCTACGGTTTGTGGTAGTAATCTTAATACC	1113
Qy	3041	TTAACCCACATGCTCAAAATCAAAATGATACATATTTCTGTAGAGACCCAGCAATACCATA	3100
Db	1114	TTAACCCACATGCTCAAAATCAAAATGATACATATTTCTGTAGAGACCCAGCAATACCATA	1173
Qy	3101	AGAAATTAATAAAAAAAAAAAAAA 3120	
Db	1174	AGAAATTAATAAAAAAAAAAAAAA 1193	
RESULT 11			
US-09-976-674-10			
; Sequence 10, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 10			
; LENGTH: 1356			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-976-674-10			
Query Match 29.7%; Score 926.4; DB 3; Length 1356;			
Best Local Similarity 99.9%; Pred. No. 2.8e-261;			
Matches 927; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	AAAGTCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG	60

Db	1	AAAGTCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTTCGCCGCTCGGTTGTCAACCGGGCGCCGCCCGAGGAAGCACTGCAACACGAGACCG	120
Db	61	CGTTTCGCCGCTCGGTTGTCAACCGGGCGCCGCCCGAGGAAGCACTGCAACACGAGACCG	120
Qy	121	GAGTGGAGGCGCGCAGCATGAAGCGCGCAGGCCCGCTCCATAGGCGCACGTTCGGGACGG	180
Db	121	GAGTGGAGGCGCGCAGCATGAAGCGCGCAGGCCCGCTCCATAGGCGCACGTTCGGGACGG	180
Qy	181	TCCGGCGGGCGCGGGGGAAGAAATGCAACATCGCAGCAGCAATGGAACAGAAACAG	240
Db	181	TCCGGCGGGCGCGGGGGAAGAAATGCAACATCGCAGCAGCAATGGAACAGAAACAG	240
Qy	241	CTGGGTGTTGAGATATTTGAAAATGCGGACTGTGAGAGAAATATTTGAATCAGAGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAAAATGCGGACTGTGAGAGAAATATTTGAATCAGAGATCGG	300
Qy	301	CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTGACGCTTAAAAAGCTGCTT	360
Db	301	CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTGACGCTTAAAAAGCTGCTT	360
Qy	361	GCCGATACAGAAAATATCATGGCTACATGATGGCTAAGCAGCACCAATGATTTTCATGTTT	420
Db	361	GCCGATACAGAAAATATCATGGCTACATGATGGCTAAGCAGCACCAATGATTTTCATGTTT	420
Qy	421	GTGAAGAGAAATGATCCAGATGGACCTCAITTCAGACAGAAATCTATTACCTTGCATGTCT	480
Db	421	GTGAAGAGAAATGATCCAGATGGACCTCAITTCAGACAGAAATCTATTACCTTGCATGTCT	480
Qy	481	GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTCCTCAAAATCTCAATAGAGCA	540
Db	481	GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTCCTCAAAATCTCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACCTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACCTGGACTAT	600
Qy	601	GGAAATGTTCTCGAGAGAAAGAACTATTAAAGAAAGAAAGCAATTCGGAACAGTCGGA	660
Db	601	GGAAATGTTCTCGAGAGAAAGAACTATTAAAGAAAGAAAGCAATTCGGAACAGTCGGA	660
Qy	661	ATTGCTTCTTACGATTATCAACAGGAAGTGGAACATTTCTGTTTCAAGCCGCTAGTGA	720
Db	661	ATTGCTTCTTACGATTATCAACAGGAAGTGGAACATTTCTGTTTCAAGCCGCTAGTGA	720
Qy	721	ATTATTCAGTAAAAAGATGGAGGCGCCAAAGGATTTACGCAACAACTTTTAAGGCCCAAT	780
Db	721	ATTATTCAGTAAAAAGATGGAGGCGCCAAAGGATTTACGCAACAACTTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAACCTAGTTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAACCTAGTTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACACAGAA	900
Db	841	GACTGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACACAGAA	900
Qy	901	GAAGGAGACTCATTATGTGCACATG 928	
Db	901	GAAGGAGACTCATTATGTGCACATG 928	
RESULT 12			
US-09-976-674-4			
; Sequence 4, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			

FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976.674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 2617
TYPE: DNA
ORGANISM: Homo sapiens
US-09-976-674-4

Query Match 27.6%; Score 861; DB 3; Length 2617;

Best Local Similarity 59.4%; Pred. No. 6.9e-242;
Matches 1499; Conservative 0; Mismatches 1020; Indels 6; Gaps 2;

Qy	316	TTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTCCGATACCAAGAAA	375
Db	80	TTCCAGGTGCAGAACGACTCGTGGACGGGCTCCGGAGCATATCCACGGCAGCCGCAAG	139
Qy	376	TATCATGCTCATATGATGGCTTAAGGCACACATGATTTCAATGTTTGTGAAGAGGAATGAT	435
Db	140	TACTCGGCGCTCATGTCAACAGCGGCCCAAGACTTCCAGTTTGTGACAGAGCGAT	199
Qy	436	CCAGATGACCTCATTTACAGACAGAAATCTATTACCTTGGCCATGTCTGGTGAGAACAGAAA	495
Db	200	GAGTCTGGGCCCTACTCCACGGCTCTACTACTGGGAATGCCATATGGCAGCCGAGAG	259
Qy	496	AATACACTGTTTATTTCTGAATTTCCAAAATCATCAATAGACAGCAGCTTTAATGCTC	555
Db	260	AATCTCCCTCTACTCTAGATTTCCCAAGAGTCCGGAAGAGGCTCTGCTGCTCCTG	319
Qy	556	TCTTGAAGAGCTCTTTTGGATCTTTTTCAGGCAACACTGGAATGATGATTTCTGA	615
Db	320	TCCTGGAAGCAGATGCTGGATCATTTCCAGGCCACCGCCACCATGGGGTCTACTCTCG	379
Qy	616	GAAGAGAACTATTAGAGAAAGAAAGCAATTTGGACAGCTCGGAATTCCTTTACGAT	675
Db	380	GAGGAGGAGCTGCTGAGGAGCGGAAACGCTCGGGGTCTTGGCATCATCTCTACGAC	439
Qy	676	TATCACCAAGGAGTGAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCACGTAAAA	735
Db	440	TTCCACAGGAGAGTGGCTCTTCTCTTCCAGGCGAGCAAGCCCTTCCACTGTGCG	499
Qy	736	GATGAGGGCCACAGGATTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGT	795
Db	500	GACGGCGCAAGACGGCTTCATGTTGTTCCCTTATGAAACCGCTGGAATCAAGACCCAG	559
Qy	796	TGTCCTCAACATACGAGTGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTT	855
Db	560	TGCTCAGGGCCCGGATGGACCCCAAAATCTGGCTCCGACCCCTCTCTCTCTCTCTC	619
Qy	856	ATACATAGCAACGATTTTGGATATCTAATCTGTAACAGAGAGAAAGAGAGCTACT	915
Db	620	ATCAATAACAGCAGCTGTGGTGGCCAAATCATCGACAGCGGAGGAGCGGGCTGACC	679
Qy	916	TATGTGCAATGAGCTAGCCAACTATGAAGAGATGTCAGATCAGCTGGAGTCGCTACC	975
Db	680	TTCTGCCCAAGGTTTATCCAATGTCTCTGATGACCCCAAGTCTCGGGGTGGGCCACC	739
Qy	976	TTTGTCTTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAA	1035
Db	740	TTCTGTATACAGGAAGAGTTCCGACCGTCTTCACTGGGTACTGGGTGGCCCAAGCTCC	799
Qy	1036	AC---AATCTCCAGTGGTGAATAATTTCTTAGAATTTCTATGAAGAAATGATGATCT	1092
Db	800	TGGGAAGTTTACAGGGGCTCAAGACGCTGGGAATCTCTGTATGAGGAATGCGATGATCC	859
Qy	1093	GAGGTGGAAATTTATCTGTTTACATCCCTATGTTGGAAACAAGAGGGGAGATTTCTTC	1152
Db	860	GAGGTGGAGGTCAATTCAGTCTCCCTCTCTCTGGCTAGAGAAAGGAAGGAGCGACTGTAT	919

Qy	1153	CGTTATCTTAAACAGGTCACAGCAAAATCCTAAAGTCATCTTTTAAGATGTGAGAAATATG	1212
Db	920	CGGTACCCAGGACAGGCAAGAAATCCCAAGATTCCTTTGAAATGGCTGAGTTCCAG	979
Qy	1213	ATTGATGCTGAAGGAGGATCATAGATGTATAGTAAGGAACTAATTTCAACCTTTTGAG	1272
Db	980	ACTGACAGCCAGGGCAAGATCGTCTCGACCCAGGAGAGGAGCTGGTGCAGCCCTTCAGC	1039
Qy	1273	ATTCTATTGAAGGAGTTGAAATATATTCGACAGAGCTGGATGAGTCTCTGAGGGAAAAAT	1332
Db	1040	TGCTGTTTCCGAAAGGTGGAGTACATCCGACGGGCGGGTGGACCCGGGATGGCAATATC	1099
Qy	1333	GCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACT	1392
Db	1100	GCTGGGCCATGTTCTTGACCGGCCCCAGCAGTGGCTCCAGCTCGTCTCTCTCCCGG	1159
Qy	1393	GAATTAATTTATCCAGTGAAGATGATGTTATGAAAGGCGAGAGCTCATTTGAGTCAGTG	1452
Db	1160	GCCTGTTTCATCCCGACACAGAAATGAGGAGCGGGCTAGCTCTGCGCAGAGCTGTC	1219
Qy	1453	CTGATTTGTGACGCGCACTAATTAATCTATGAAGAAACAACAGACATCTGGATTAATATC	1512
Db	1220	CCAGGAATGTCCAGCGGTATGTGTGTACGAGGAGGTCAACAACGCTCTGGATCAATGTT	1279
Qy	1513	CATGACATCTTTCATGTTTTCCTCCAAAGTCACGAGAG---GAAATTGAGTTTATTTT	1569
Db	1280	CATGACATCTTTCATGTTTTCCTCCCAATCAGAGGAGGAGCGAGCTCTGCTTTCTCCGC	1339
Qy	1570	GCCTCTGAATGAAACAGGTTTCCTCATTTATTAATAAATTAATCATATTTTAAAGGAA	1629
Db	1340	GCAATGAATGAAGACGGCTTCTGCCATTTGTACAAGATCAACGGCGTTTAAATATCC	1399
Qy	1630	AGCAATATAAAGCATCCAGTGGTGGCTGCTCCAAAGTGAATTTCAAGTGTCTTATC	1689
Db	1400	CAGGGCTACGATTTGAGTGAGCCCTTCAGCCCGGGGAGATGAATTTAGTGCCTCAT	1459
Qy	1690	AAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGCCCGCATGGAATCTAAT	1749
Db	1460	AAGGAGAGATTTGCTCTGACAGCGGTGAATGGGAGGTTTGGCGAGGACGGCTCCAG	1519
Qy	1750	ATCAAGTTGATGAAGTCAAGAGGTGGTATATTTTGAAGGACCAAGAGATCCCTTTTA	1809
Db	1520	ATCTGGGTCAATGAGGAGACCAAGCTGTGTACTTCCAGGCGACCAAGGACACGCGGTG	1579
Qy	1810	GAGCATCACTGTAGTAGTCAATAGTAAATCTCTGAGAGGTGACAGAGCTCACTGAC	1869
Db	1580	GAGCACCACTCTACGTGTGCTAGTATGAGGCGGCGCGGAGATCGTACGCTCACCACG	1639
Qy	1870	CGTGGCTACTACATTTCTTGCTGATCAGTCAGCACTGTGACTTCTTTATAGTAAGTAT	1929
Db	1640	CCCGGCTTCTCCATAGCTGCTCCATGAGCCAGAACTTCGACATGTTGCTGAGCCACTAC	1699
Qy	1930	AGTAAACAGAGAAATCCACATGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGAC	1989
Db	1700	AGCAGGTGAGCAGCGCCCTGCGTGCACGTCTTCAAGCTGAGGGGCGCCGACGACGAC	1759
Qy	1990	CCAATTTGCAAAACAAAGGAATTTTGGGCCCACTTTTGGATTTGAGTTCAGCAGGTCCT	2049
Db	1760	CCCTGCAACAGGAGCCCGCTTCTGGGCTAGCATGATGAGGAGCAGCAGCTGCCCCCG	1819
Qy	2050	GACTATATCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTTGATGGG	2109
Db	1820	GAATGTTCTCCAGAGATCTTCCATTTCCACACGCGCTCGGATGTGCGGCTCTACGCG	1879
Qy	2110	ATGCTCTACAGCCTCATGATCTACAGCTGGAAGAAATATCTTACTGTGCTGTTCTATA	2169
Db	1880	ATGATCTACAGCCCCCAGCTTGCAGCCAGGAGGAGAGACCCCAACCGCTCTCTTTGTA	1939
Qy	2170	TATGTTGGTCTCAGGTGAGTGTGTTGAATATTCGTTTAAAGGAGTCAAGTATTTCCGC	2229
Db	1940	TATGAGGCGCCCGAGTGCAGTGTGTAATTAATCTCTTCAAGGCAATCAAGTACTTGGG	1999
Qy	2230	TTGAATACCTAGCCTCTCTAGTGTATGTTGTTGATAGACAAACAGGGGATCTCTGT	2289

Db 2000 CTCACACACTGGCTCCCTGGGCTAGCGCGTGTGTGATGACGGCAGGGGCTCCGT 2059
Qy 2290 CACGAGGCTTAATTTGAAGGCGCTTTAAATATATAATGGGTCAATAGAAATGAC 2349
Db 2060 CAGGAGGCTTCGGTTGAAAGGCGCTCGAAACCAATGGGCCAGGTGGAGTCAG 2119
Qy 2350 GATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATGACTAGATCGT 2409
Db 2120 GACAGGTGGAGGCGCTGCGAGTTCGTGGCCGAGAGATATGGCTTCATCGACTGAGCCGA 2179
Qy 2410 GTGGGCATPCACGGCTGGTCTATGAGAGATACCTCTCCCTGANTGCCATTAATGAGAGG 2469
Db 2180 GTTGCCATPCCATGGCTGGTCTACGGGGGCTTCTCTCGCTCATGGGGCTAATCCACAAG 2239
Qy 2470 TCAGATATCTTCAGGGTGTCTATGTCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGAT 2529
Db 2240 CCCAGGTGTTCAAGGTGGCCATCGCGGGTGGCCCCGGTCAACCGTCTGGATGGCCTACGAC 2299
Qy 2530 ACAGGATACACGGAACTGATATATGGGTCACTCTGACGAGAAATGAACAGGGCTATTACTTA 2589
Db 2300 ACAGGGTACATGAGCGCTACATGGACGCTCCCTGAGAAACAACAGCACGGCTATGAGGCG 2359
Qy 2590 GGATCTGTGGCCATGCAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTTA 2649
Db 2360 GGTTCGTGGCCCTGCAGCTGGAGAGCTGCCCAATGAGCCCAACCGCTTGTATCTCTC 2419
Qy 2650 CATGTTTCTCGATGAGATGCTCATTTTGCATACCAATGATATATGAGTTTCTTA 2709
Db 2420 CACGGCTTCTTGAAGAAACGTGCACTTTTCCACAAAATTCTCTGCTCTCCCACTG 2479
Qy 2710 GTGAGGGTGAAGAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAGATTAAGA 2769
Db 2480 ATCCGAGCAGGAAACCTTACAGCTCCAGATCTACCCCAACGAGAGACACAGTATTCG 2539
Qy 2770 GTTCTGTAATCGGAGAACATATGAACTGCATCTTTTGCATCACTTCAAGAAACCTT 2829
Db 2540 TGCCCCGAGTGGGCGAGCACTATGAAGTCACGTTGCTGCACCTTCTACAGGAATACCTC 2599
Qy 2830 GGATC 2834
Db 2600 TGAGC 2604

RESULT 13
; US-09-976-674-28
; Sequence 28, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70869
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 4219
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-28

Query Match 27.6%; Score 861; DB 3; Length 4219;
Best Local Similarity 59.4%; Pred. No. 9.3e-242;
Matches 1499; Conservative 0; Mismatches 1020; Indels 6; Gaps 2;
Qy 316 TTTTATGTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCAAAAA 375

Db 436 TTCCAGGTGCAGAAAGCACTCGTGGGACGGGCTCCGAGCATCATCCACGGCAGCCGCAAG 495
Qy 376 TATCATGGCTACATGATGGCTAAGGCACCAATGATTTTGTGTAAGAGGAATGAT 435
Db 496 TACTCGGCGCTCATTTGTCAACAGGCGCCCGACGACTTCAGTTTGTGCAGAAAGCGGAT 555
Qy 436 CCAGATGGACCTCATTTACAGACAGAAATCTATTACCTTCGCATGTCTGGTGAGAAACAGAGAA 495
Db 556 GAGTCTGGGCCCCACTCCACCGCTCTACTACCTGGGAATGCCATATGACGCCGAGAG 615
Qy 496 AATACTACTTTTATCTGGAATTTCCCAAACTATCAATAGACGACAGCTCTTAATGCTTC 555
Db 616 AACTCCTCTCTACTCTGAGTTCCCAAGAGTTCGGAAGAGGCTCTGCTGCTCTG 675
Qy 556 TCTTGAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTTATCTCGA 615
Db 676 TCTTGAAGCAGATGCTGGATCATTTCCAGGCCACGCCCAACCATGGGTCTACTCTCGG 735
Qy 616 GAAGAAGAACTATTAAAGAAAGAAAGCAAGTTCGAAACAGTTCGGAATGTTCTTACGAT 675
Db 736 GAGGAGGAGCTGCTGAGGGAGCGGAAGCGCTGGGGTCTTTCGGCATCACTCTCTACGAC 795
Qy 676 TATCAACGAAGAGTGGAAACATTTCTGTTCAGCCCGGTAGTGGAAATTTATACGTAATAA 735
Db 796 TTCCACAGCAGAGTGGCTCTTCTCTTCCAGGCCAGCAACAGCTCTTCCACTGCGCG 855
Qy 736 GATGAGGGCCACAAAGGATTTACGCAACAACCTTTAAAGGCCCAATCTAGTGGAAACTAGT 795
Db 856 GACGGCGCAAGAACCGCTTCATGGTGTCCCTATGAACCGCTGGAAATCAAGACCCAG 915
Qy 796 TGTCCCAACATACGATGGATGATCCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTT 855
Db 916 TGCTCAGGGCCCGGATGGACCCCAAAATCTGCCCTGCGGACCTGCTCTCTCTCTCTC 975
Qy 856 ATACATAGCAAGATTTTGGATCTTAAAGTCTTAACGATGTAACGAGAGAAAGAGACTCACT 915
Db 976 ATCAATACAGCGACTGTGGTGGCCAACTCAGACAGCGAGAGCGGGCTGACC 1035
Qy 916 TATGTGCACATGAGCTAGCCAAACATGGAAGAAGATGCGAGATCAGCTGAGTGGCTACC 975
Db 1036 TTTCTGCCAACAGGTTTATCCAATGTCCTGATGACCCCAAGTCTGCGGGTGGGCCACC 1095
Qy 976 TTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAA 1035
Db 1096 TTTCTGCATACAGAAAGAGTTCGACCGCTTCACTGGGTACTGGTGGTGGCCCAAGCTCC 1155
Qy 1036 AC---AATCTCCAGTGGTGGTAAATTTCTTAAAGTTCTATATGAAGAAATATGATGATCT 1092
Db 1156 TGGGAAGGTTACAGAGGGCTCAAGACGCTGCGAATCTCTGTATGAGGAAGTCGATGAGTCC 1215
Qy 1093 GAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAACAAGAGGAGGCGAGATTCATTTC 1152
Db 1216 GAGGTGAGGTCAATTCAGTCCCTCTCTCGCTAGAAAGAAAGACGGAATCGTAT 1275
Qy 1153 CGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATATG 1212
Db 1276 CGGTACCCAGGACAGGACGCAAGATCCCAAGATTGCTTGAACCTGGCTGAGTTCCAG 1335
Qy 1213 ATTGATGTGAAGAGGATCATAGATGTCTAGATAGGAGCAATTAATCACTTTTGGAG 1272
Db 1336 ACTGACAGCCAGGCAAGATGCTCTGACCAAGAGAGAGAGCTGTTGCGAGCCCTTCAGC 1395
Qy 1273 ATTCTATTGAAGAGTTGAATATATTCGACAGCTGATGAGTCTCTGAGGGAATAT 1332
Db 1396 TCGCTGTTCCCAAGGTGGAGTACATCGCAGGGCGGGTGGACCCGGGATGGCAATATC 1455
Qy 1333 GCTTGTCCATCTACTAGATCGCTCCAGACTCGCTPACAGATAGTGTGATGATCTCACT 1392
Db 1456 GCCTGGGCATGTTCTTGGACCGGCCCCAGCAGTGGCTCCAGCTGCTCTCTCTCCCCCG 1515
Qy 1393 GAATTTATTTATCCAGTGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAAGT 1452
Db 1516 GCCTGTTTATCCGAGCACAGAAATGAGGAGCAGCGGCTAGCCTCTGCGAGAGCTGTC 1575

Db	736	GAGGAGAGCTGCTGAGGAGCGGAAACGCTGGGGTCTTCGGCATCACCTCTCAACG	1816	RAGGAAGAGATTGCTCTGACCAGCGGTGAATGGAGGTTTGGCGGAGCAGCGCTCCAAG
Qy	735	TATCACCAAGGAAGTGGAACATTTCTGTTTCAAGCGGAGTGGAAATTTATCAGTAAAA	1750	ATCCAAGTTGATGAAGTCTGATATATTTTGAAGGACCAAAAGACTCCCTCTTTA
Db	735	TTCCACAGCGAGAGTGGCTTCTCTCTTCCAGGCGCAGCAACAGGCTCTTCCACTGCCGC	1876	ATCTGGGTCAATGAGGAGACCAAGCTGGTCTACTTCCAGGACCAAGGACACGCGCTG
Qy	795	GATGAGGGCCACAGGATTTACGCAACAACTTTTAAGGCCCAATCTCTAGTGGAACTAGT	1810	GAGCATCACCTGTAGTACGTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGAC
Db	915	GACGGCGCAAGAACCGCTTCATGGTGTCCCTTATGAAACCGCTGGAAATCAAGACCCAG	1936	GAGCACCACTCTACGTGGTCACTATGAGCGCGCGCGAGATCGTACGCCTCACCAAG
Qy	855	TGTCTCAACATACCGATGGATGCCAAATATGCCCCGCTGATCCAGACTGGATGCTTTT	1870	CGTGCTACTCACATTTCTGCTGATCAGTCAGCATGCTGACTCTCTTATAGTAGTAT
Db	975	TGCTCAGGCGCCGATGAGACCCCAAAATCTGCCCTGCCGACCTGCTCTTCTCTCTTC	1996	CCCGGCTTCTCCCATAGCTCTCATGAGCGCAACTTCGACATGTTCTGCACCACTAC
Qy	915	ATACATAGCAACAGATTTTGGATATCTAACATCGTAAACAGAGAAAGAGAGACTCACT	1930	AGTAAACCAAGAAATCCACACTGTGTCTCTTTTCAAGCTATCAAGTCTCTGAAGTAC
Db	1035	ATCAATAACAGGACCTGTGGTGGCCAAACATCGACAGCGAGGAGCGCGCTGACC	2056	AGCAGCGTGAGCAGCGCGCTCTGCTGACAGTCTACAAAGTGTAGGAGCGCCGACGAC
Qy	975	TATGTGCAATGAGCTAGCCACATCGAAGNAGATGCCAGATCAGCTGGAGTCCGTACC	1990	CCAACTTGCAAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTC
Db	1036	TTCTGCCACCAAGTTTATCCATGTCTCGATGACCCCAAGTCTGCGGTGTGGCCACC	2116	CCCTGTGCAACAGCAGCCCGCTTCTGGGCTAGCATGATGAGGCGAGCAGCTGCCCCCG
Qy	1035	TTTGTCTCBAAGAAATTTGATAGATATTTCTGGCTATTTGGGTGTCCTCAAAAGCTGAA	2050	GACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGG
Db	1096	TTGTCATACAGGAGAGTTCGACCGCTTCACTGGGTACTGGTGGTGGCCCAACAGCTCC	2176	GATTAATGTTCTCCAGAGATCTTCCATTTCCACACGCGCTCGATGTGGGGCTCTACGCG
Qy	1092	AC--AATCCCAAGTGGTGTAAATTTCTTAGAATTTCTATATGAGAAATGATGAATCT	2110	ATGCTCTACAAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTATA
Db	1156	TGGGAAGGTTTCAAGAGGCGCTCAAGACGCTCGCAATCTGTATGAGNAGTCGATGAGTCC	2236	ATGATCTACAGCCCGCTTGCAGCGAGGAAAGAACACCCACCGTCTCTTTGTA
Qy	1152	GAGGTGGAATTTATCATGTTTACATCCCTTATGTTGGAACAAAGAGGCGAGATTCATTC	2170	TATGCTGTCTCAGGTGCTGAGTGTGATGTTGGTGTAGTGTAGATAGACACAGGGGATCTGT
Db	1216	GAGGTGAGAGTCATTCAGCTCCCTCTCTCGCTAGAAAGAAAGGAGAGCGAGCTGGAT	2296	TATGAGGCGCCCGAGTGCAGCTGTGTGATGATGCGCGTGTGTGATTTGACGCGAGGGCTCTGT
Qy	1212	CGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGTGTGAGAAATAATG	2230	TTGAATACCTTAGCTCTCTAGGTTATGTTGGTGTAGTGTAGATAGACACAGGGGATCTGT
Db	1335	CGGTACCCAGGACAGGCAAGAAATCCCAAGATTTGCCCTTGAATCGGCTGAGTCCAG	2356	CTCAACACACTGGCTCTCTGGCTACGCGTGTGTGATTTGACGCGAGGGCTCTGT
Qy	1272	ATTGATGCTGAAGGAAGATCATAGATGTATAGATAAGGAATTAATTCACACTTTTGAG	2290	CACGAGGCGCTTAAATTTGAAGGCGCTTTAAATATAAATGAGTCAAGTAAATGAC
Db	1336	ACTGACAGCGGCGAAGATCGTCTCGACCCAGGAGAGGAGCTGGTGCAGCGCTTCAGC	2416	CAGCAGGCGCTTCGGTTCGAAGGGCGCTGAAACCAAAATGGCGAGTGGAGATCGAG
Qy	1332	ATTCATTTGAAGGATTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAAT	2350	GATCAGGTGGAAAGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTTGATAGATCGT
Db	1455	TCGCTGTTCCGAGGTGAGTACATCGCCAGGCGGGTGGACCCCGGATGGCAATAC	2476	GACCAGGTGGAGGCGCTGAGTTCGTGGCCGAGAAAGTATGGCTTCATCGACCTGAGCCGA
Qy	1392	GCTTGGTCACTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACT	2410	GTGGGCATCCAGCGCTGCTCTTATGGAGGATACCTCTCCCTGATGGCATTTAATGCGAGG
Db	1456	GCTGGCCATGTTCTTGACCGGCCCCAGCAGTGGCTCCAGCTCTCTCTCTCCCGCG	2536	GTTCGCATCCATGGCTGCTTACGCGGCGCTTCTCTCGCTCATGGGGCTTAATCCACAAG
Qy	1393	GAAATTTATCCCAAGTAAAGATGATGTTATGGAAGGCGAGAGACTCAATTGAGTCAGTG	2470	TCAGATATCTTTCAGGGTGTGCTATTGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGAT
Db	1516	GCCCTGTTATCCGAGGACAGAGATGAGGAGCGCGGTAGCCTCTGCGAGAGCTGTC	2596	CCCCAGTCTTCAAGGTGGCCATCGCGGCTGCCCGGTCAACCGTCTGGATGGCTACGAC
Qy	1453	CCTGATTTCTGAGCCCACTAATTTATGAGAAACAAACAGACATCTGGATTAATATC	2530	ACAGGATACACGGAACGTTATATGGTCACTGACAGAAATGAACAGGGGCTATTACTTA
Db	1635	CCAGGAATGTCCAGCCGATGTGGTGTACAGAGAGTCCAAACGCTGGATCAATGTT	2656	ACAGGTACACTGAGGCGTACATGGACGTCCTCGAAGAACACAGCAGCGCTATGAGCG
Qy	1569	CATGACATCTTTCATGTTTTTCCCCAAAGTCAAGAG--GAAATGAGTTTATTTTT	2590	GGATCTGGCCATCGAAGCAAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTA
Db	1636	CATGACATCTTCTATCTCTCCCTCCCAATCAGAGGAGAGAGCGAGCTCTGTTCTCCGC	2716	GGTTCGTGGCTCTGCACTGGAGAACTGCGCAATGAGCCCAACCGCTGCTTATCCTC
Qy	1629	GCTCTGAAATGCAAAACAGGTTTCGTCATTTATACAAAATACATCTATTTTAAAGGAA	2650	CATGTTTCTCGATGAGATGTCATTTTTCGACATACCACTATTAATCTAGTGTTTTAA
Db	1696	GCCAAATGAATGCAAGACCGGCTTCTGCCATTTGTACAAAGTCAACCGCGTTTTTAAATCC	2776	CAGGGTCTCTGAGCAAAACGTGCTTTTTTCCACAAACTTCTCTCGTCCCACTG
Qy	1630	AGCAATATAAACAGATCCAGTGGTGGCTGCTCTCAAGATGATTTCAAGTGTCTATC	2710	GTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTGAGGAGACACAGCATAGA
Db	1756	CAGGGCTACGATTTGAGTGAGCCCTTCAGCCCGGGGAAGATGAATTAAGTCCCAAT	2836	ATCCGAGCAGGGAACCTTACAGCTCCAGATCTACCCCAACAGAGAGACACATGTTTCG
Qy	1690	AAAGGAGATAGCAATACAGTGGTGAATGGGAAGTTCTTGGCGCGCATGGATCTAAT	2770	GTTCTGAAATCGGGAGAACATTTGAACTGGCATTTTTTGGCTACTCTTCAAGAAACGTT
			2896	TGCCCGGAGTCCGGCGAGCACTATGAAGTCAGTTGCTGCTGCTTCTACAGGAATACCTC

Qy 2830 GGATC 2834
Db 2956 TGAGC 2960

RESULT 15
US-09-976-674-18
; Sequence 18, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-18

Query Match 25.7%; Score 801.4; DB 3; Length 832;
Best Local Similarity 99.5%; Pred. NO. 1.1e-224;
Matches 799; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGTGTAAAGCTCGAGGCCAAGCGCGTCTACTCGCGCGCTCTTCTTAGTGCCG 60
Db 1 AAGTGTAAAGCTCGAGGCCAAGCGCGTCTACTGCGCGCGCTCTTCTTAGTGCCG 60

Qy 61 CGTTCCGCCCTGGGTGTCTACCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCG 120
Db 61 CGTTCCGCCCTGGGTGTCTACCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCG 120

Qy 121 GAGTGAGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCATCTCGGACCG 180
Db 121 GAGTGAGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCATCTCGGACCG 180

Qy 181 TCCGGCGCGCGCGCGGGAAGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
Db 181 TCCGGCGCGCGCGCGGGAAGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240

Qy 241 CTGGGTGTTGAGATATTTGAAACTCGCGACTGTGAGGAGAAATTTGAATCAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTCGCGACTGTGAGGAGAAATTTGAATCAGGATCGG 300

Qy 301 CTTAAATCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCTAGCTTAAAAAGCTGCTT 360
Db 301 CTTAAATCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCTAGCTTAAAAAGCTGCTT 360

Qy 361 GCCGATACAGAAATATCATGCTACATGATGCTAAGGCACCATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAATATCATGCTACATGATGCTAAGGCACCATGATTTTCATGTTT 420

Qy 421 GTGAGAGGAATATCCAGATGACCTCATTCAGACAGATCTATTACCTTGGCATGTCT 480
Db 421 GTGAGAGGAATATCCAGATGACCTCATTCAGACAGATCTATTACCTTGGCATGTCT 480

Qy 481 GGTGAGACAGAGAAATACACTGTTTTTATTCGAAATTCCTGAAATCTATCAATAGAGCA 540
Db 481 GGTGAGACAGAGAAATACACTGTTTTTATTCGAAATTCCTGAAATCTATCAATAGAGCA 540

Qy 541 GCAGTCTTAATGCTCTCTTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600

Search completed: April 15, 2006, 00:11:57
Job time : 542 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 22:13:06 ; Search time 2408 Seconds
(without alignments)
10714.476 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 agtgcctaaagcctccgagg.....agaattactataaaaaaaa 3120

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	%	DB	ID	Description
1	3120	100.0	3120	7	US-10-415-122-5	Sequence 5, Appli
2	3120	100.0	3120	8	US-10-825-632-2	Sequence 2, Appli
3	3106.4	99.6	4829	3	US-09-976-674-12	Sequence 12, Appl
4	3106.4	99.6	4829	3	US-10-982-512-12	Sequence 12, Appl
5	3105.2	99.5	3143	6	US-10-170-789-37	Sequence 37, Appl
6	3101.8	99.4	3106	7	US-10-311-035-30	Sequence 30, Appl
7	2824.4	90.5	4685	3	US-09-976-674-22	Sequence 22, Appl
8	2824.4	90.5	4685	3	US-10-982-512-22	Sequence 20, Appl
9	2806.4	89.9	4676	3	US-09-976-674-20	Sequence 20, Appl
10	2806.4	89.9	4676	3	US-10-982-512-20	Sequence 20, Appl
11	2645.8	84.8	2649	6	US-10-170-789-39	Sequence 39, Appl
12	2644.4	84.8	2671	3	US-09-976-674-2	Sequence 2, Appli
13	2644.4	84.8	2671	9	US-10-982-512-2	Sequence 2, Appli
14	2644.2	84.7	2649	6	US-10-054-776-1	Sequence 1, Appli
15	2508.4	80.4	4523	3	US-09-956-157-2177	Sequence 2177, Ap
16	2508.4	80.1	4523	3	US-09-976-674-8	Sequence 8, Appli
17	2500.4	80.1	4523	9	US-10-982-512-8	Sequence 8, Appli
18	2208	70.8	2510	7	US-10-275-505-16	Sequence 16, Appl
19	2208	70.8	2510	10	US-11-140-224-16	Sequence 16, Appl
20	2056.4	65.9	4309	3	US-09-976-674-14	Sequence 14, Appl
21	2056.4	65.9	4309	3	US-10-982-512-14	Sequence 14, Appl
22	1347	43.2	1669	8	US-10-825-632-6	Sequence 6, Appli
23	1036	33.2	1197	8	US-10-825-632-4	Sequence 4, Appli

24	926.4	29.7	1356	3	US-09-976-674-10	Sequence 10, Appl
25	926.4	29.7	1356	7	US-10-982-512-10	Sequence 10, Appl
26	883.8	28.3	3287	9	US-10-415-122-3	Sequence 3, Appli
27	862.6	27.6	2660	7	US-10-072-012-223	Sequence 223, App
28	862.6	27.6	2660	7	US-10-072-012-225	Sequence 225, App
29	861	27.6	2617	3	US-09-976-674-4	Sequence 4, Appli
30	861	27.6	2617	9	US-10-982-512-4	Sequence 4, Appli
31	861	27.6	3716	9	US-10-433-757-30	Sequence 30, Appl
32	861	27.6	4219	3	US-09-976-674-28	Sequence 28, Appl
33	861	27.6	4219	9	US-10-982-512-28	Sequence 28, Appl
34	861	27.6	4302	3	US-09-976-674-24	Sequence 24, Appl
35	861	27.6	4302	3	US-10-982-512-24	Sequence 24, Appl
36	857.8	27.5	3000	7	US-10-415-122-1	Sequence 1, Appli
37	854	27.4	2495	7	US-10-415-122-8	Sequence 8, Appli
38	801.4	25.7	832	3	US-09-976-674-18	Sequence 18, Appl
39	801.4	25.7	832	9	US-10-982-512-18	Sequence 18, Appl
40	798.6	25.6	4180	3	US-09-976-674-36	Sequence 36, Appl
41	798.6	25.6	4180	9	US-10-982-512-36	Sequence 36, Appl
42	798.6	25.6	4263	3	US-09-976-674-34	Sequence 34, Appl
43	798.6	25.6	4263	9	US-10-982-512-34	Sequence 34, Appl
44	791.2	25.4	1083	8	US-10-825-632-8	Sequence 8, Appli
45	787.6	25.2	925	6	US-10-264-237-710	Sequence 710, App

ALIGNMENTS

RESULT 1
US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: P15217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-415-122-5

Query Match 100.0%; Score 3120; DB 7; Length 3120;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AAGTGTAAAGCCTCCGAGCCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
DB	1	AAGTGTAAAGCCTCCGAGCCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
QY	61	CGTTGCGCGCGCTGGTGTTCACCGCGCGCGCGCGAGGAGCCACTCAACAGGACCG	120
DB	61	CGTTGCGCGCGCTGGTGTTCACCGCGCGCGCGAGGAGCCACTCAACAGGACCG	120
QY	121	GAGTGGAGGCGCGGACGATGAGCGCGCGCGCGCTCCATAGCGACGTCGGGACCG	180
DB	121	GAGTGGAGGCGCGGACGATGAGCGCGCGCGCGCTCCATAGCGACGTCGGGACCG	180
QY	181	TCCGCGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGACAG	240
DB	181	TCCGCGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGACAG	240
QY	241	CTGGGTGTTGAGATATTGAAACTCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	300
DB	241	CTGGGTGTTGAGATATTGAAACTCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	300
QY	301	CCTAAATGGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGCTGCTT	360
DB	301	CCTAAATGGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGCTGCTT	360


```
Db 2521 TTCTATGATACAGGATACACGGAACGTTTATATGGGTCAACCTGACAGAAATGAACAGGGC 2580
Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Db 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Qy 2641 CTGCTCTTACATGGTTTCTCGGATGAGAAATGTCATTTTGCATACCAATGATATCTAGT 2700
Db 2641 CTGCTCTTACATGGTTTCTCGGATGAGAAATGTCATTTTGCATACCAATGATATCTAGT 2700
Qy 2701 AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACAC 2760
Db 2701 AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACAC 2760
Qy 2761 AGCATAAGAGTTCTCTGAATCGGAGAACATTTATGAATGCACTCTTTTGCACCTTCAA 2820
Db 2761 AGCATAAGAGTTCTCTGAATCGGAGAACATTTATGAATGCACTCTTTTGCACCTTCAA 2820
Qy 2821 GAAAACTTGGATACGTTATTTGCTGCTCTAAAGTGATATAATTTTGAACCTGTGAGAAC 2880
Db 2821 GAAAACTTGGATACGTTATTTGCTGCTCTAAAGTGATATAATTTTGAACCTGTGAGAAC 2880
Qy 2881 TCTCTGGTATACACCTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA 2940
Db 2881 TCTCTGGTATACACCTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA 2940
Qy 2941 ATTGATCATCATTTTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
Db 2941 ATTGATCATCATTTTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
Qy 3001 CCATGAGGGTCTACGGTTTGTGGTAGTAATCTTAATCACTTAACCCCATGCTCAAAA 3060
Db 3001 CCATGAGGGTCTACGGTTTGTGGTAGTAATCTTAATCACTTAACCCCATGCTCAAAA 3060
Qy 3061 TCAATGATACATATTTCTGAGAGACCCAGCAATACCATAAGAAATTAATAAATAAATAA 3120
Db 3061 TCAATGATACATATTTCTGAGAGACCCAGCAATACCATAAGAAATTAATAAATAAATAA 3120
```

RESULT 2

```
US-10-825-632-2
; Sequence 2, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSE-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-2
```

```
Query Match 100.0%; Score 3120; DB 8; Length 3120;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGTGCTAAGCCTCCGAGGCCAAGCGCGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 AAGTGCTAAGCCTCCGAGGCCAAGCGCGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
```

```
Qy 61 CGTTCCGCCCTGGGTTGTCAACCGCGCGCGGAGGAGCCACTGCAACAGGACCG 120
Db 61 CGTTCCGCCCTGGGTTGTCAACCGCGCGCGGAGGAGCCACTGCAACAGGACCG 120
Qy 121 GAGTGAGCGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGGACGG 180
Db 121 GAGTGAGCGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGGACGG 180
Qy 181 TCCGGCGCGGCGCGGGGAGAAAATGCAACATGCGCAGCAGCAATGGAAAACAGAACAG 240
Db 181 TCCGGCGCGGCGCGGGGAGAAAATGCAACATGCGCAGCAGCAATGGAAAACAGAACAG 240
Qy 241 CTGGGTGTTGAGATATTTGAAAATCTCGGACTGTGAGGAGATATTTGAATCAAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAAATCTCGGACTGTGAGGAGATATTTGAATCAAGGATCGG 300
Qy 301 CCTAAATTCGAGCCTTTTTTATGTTGAGCGGTATTTCTCGAGTCACTGCTTAAAGCTGCTT 360
Db 301 CCTAAATTCGAGCCTTTTTTATGTTGAGCGGTATTTCTCGAGTCACTGCTTAAAGCTGCTT 360
Qy 361 GCCGATACAGAAAATATCATGGCTACATGATGGCTAAGGCCACACATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAAATATCATGGCTACATGATGGCTAAGGCCACACATGATTTTCATGTTT 420
Qy 421 GTGAAGGAAATGATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCT 480
Db 421 GTGAAGGAAATGATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCT 480
Qy 481 GGTGAGAACAGAGAAAATACACTGTTTTTATCTGAAATTCCTCAAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAAATACACTGTTTTTATCTGAAATTCCTCAAACTATCAATAGAGCA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
Qy 601 GGAATGTAATTCGAGAGAGAAACTATTAAGAGAGAAAGAAACGATTTGGAAACAGTCGGA 660
Db 601 GGAATGTAATTCGAGAGAGAAACTATTAAGAGAGAAAGAAACGATTTGGAAACAGTCGGA 660
Qy 661 ATTGCTCTTTCATGATATCAACAGAGAGTGGACATTTCTGTTTCAAGCGGTAGTGGGA 720
Db 661 ATTGCTCTTTCATGATATCAACAGAGAGTGGACATTTCTGTTTCAAGCGGTAGTGGGA 720
Qy 721 ATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGCAACAACTTTTAAAGCCCAAT 780
Db 721 ATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGCAACAACTTTTAAAGCCCAAT 780
Qy 781 CTAGTGGAAAATAGTTGTCTCCCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAAAATAGTTGTCTCCCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA 840
Qy 841 GACTGGATGCTTTTATACATAGAACAGATATTGGATATCTAACATGTAACACAGAGAA 900
Db 841 GACTGGATGCTTTTATACATAGAACAGATATTGGATATCTAACATGTAACACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTCACAAATGAGCTAGCACAACATGGAGAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTCACAAATGAGCTAGCACAACATGGAGAGAGATGCCAGATCA 960
Qy 961 GCTGGAGTCCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAAGCTGAAAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAAGCTGAAAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATCAATCTGAGGTGGAATTTATTCATGTTTACATCCCTATGTTGGAACAGGAGG 1140
Db 1081 AATGATCAATCTGAGGTGGAATTTATTCATGTTTACATCCCTATGTTGGAACAGGAGG 1140
```

1141 GCAGATTCAATCCGTTATCTTAAACAGGTAAGCAATCCCTAAAGTCACTTTTAAAGATG 1200
Db GCAGATTCAATCCGTTATCTTAAACAGGTAAGCAATCCCTAAAGTCACTTTTAAAGATG 1200
Qy TCAGAAATAATGATGTGCTGAAAGGAGATCATAGATGTCTATAGATTAAGGAATCAATTT 1260
Db TCAGAAATAATGATGTGCTGAAAGGAGATCATAGATGTCTATAGATTAAGGAATCAATTT 1260
Qy CAACCTTTTGAGATTCTATTTGAAAGGAGTTGAATATATTTGCCAGAGCTGGATGACTCCT 1320
Db CAACCTTTTGAGATTCTATTTGAAAGGAGTTGAATATATTTGCCAGAGCTGGATGACTCCT 1320
Qy GAGGAAAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG 1380
Db GAGGAAAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG 1380
Qy TTGATCTCACCTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGAGCAGAGATCT 1440
Db TTGATCTCACCTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGAGCAGAGATCT 1440
Qy ATTTGAGTCAGTGCCTGATCTCTGACGCCACTAATATCTATGAAGAAAACACAGATC 1500
Db ATTTGAGTCAGTGCCTGATCTCTGACGCCACTAATATCTATGAAGAAAACACAGATC 1500
Qy TGGATAAATATCCATGATCTTTCATGTTTCCCAAGTCAAGAGGAGAAATTCAG 1560
Db TGGATAAATATCCATGATCTTTCATGTTTCCCAAGTCAAGAGGAGAAATTCAG 1560
Qy TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCAATTTATACAAAATACATCTATT 1620
Db TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCAATTTATACAAAATACATCTATT 1620
Qy TTTAAAGGAAGCAATATAAACGATCCAGTGGTGGCTGCTCTCCAAAGTGAATTCAG 1680
Db TTTAAAGGAAGCAATATAAACGATCCAGTGGTGGCTGCTCTCCAAAGTGAATTCAG 1680
Qy TGTCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db TGTCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Qy GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Qy TCCCTTTAGAGCATCACTCTAGTAGTACGTTACGTTAATCCTGGAGAGGTGACAAGG 1860
Db TCCCTTTAGAGCATCACTCTAGTAGTACGTTAATCCTGGAGAGGTGACAAGG 1860
Qy CTGACTGACCGTGGTACTACATCTTCTGCGATCAGTCAGTCAGCTTCTTTATA 1920
Db CTGACTGACCGTGGTACTACATCTTCTGCGATCAGTCAGTCAGCTTCTTTATA 1920
Qy AGTAAGTATAGTAACAGAGATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTT 1980
Db AGTAAGTATAGTAACAGAGATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTT 1980
Qy GAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGAGT 2040
Db GAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGAGT 2040
Qy CCTCTTCTGACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTACA 2100
Db CCTCTTCTGACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTACA 2100
Qy TTGTATGGGATCTCTACAAGCCTCATGATCTACAGCTCGGAAAGAAATCTCTACTGTG 2160
Db TTGTATGGGATCTCTACAAGCCTCATGATCTACAGCTCGGAAAGAAATCTCTACTGTG 2160
Qy CTGTTTCATATGGTGGTCTTCAGTGCAGTTGGTGAATATCGTTTAAAGGAGTCAAG 2220
Db CTGTTTCATATGGTGGTCTTCAGTGCAGTTGGTGAATATCGTTTAAAGGAGTCAAG 2220
Qy TATTTCCGCTTGAATACCTAGCCTCTAGGTTATGTGGTTAGTGATAGACACAGG 2280

2221 TATTTCCGCTTGAATACCTAGCCTCTAGGTTATGTGGTTAGTGATAGACACAGG 2280
Qy GGATCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
Db GGATCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
Qy GAAATGACGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
Db GAAATGACGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
Qy TTAGATCGGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTTA 2460
Db TTAGATCGGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTTA 2460
Qy ATGCAAGAGTCAAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATC 2520
Db ATGCAAGAGTCAAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATC 2520
Qy TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACAGAAATGAACAGGGC 2580
Db TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACAGAAATGAACAGGGC 2580
Qy TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Db TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Qy CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCACATACAGTATATTACTG 2700
Db CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCACATACAGTATATTACTG 2700
Qy AGTTTTTAGTAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACAC 2760
Db AGTTTTTAGTAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACAC 2760
Qy AGCATAAGAGTTTCTGAAATCGGAGAACATPATGAACTGTGATCTTTTGCACTACCTTCAA 2820
Db AGCATAAGAGTTTCTGAAATCGGAGAACATPATGAACTGTGATCTTTTGCACTACCTTCAA 2820
Qy GAAACCTTTGGATCAGTATTGCTGCTCTAAAGTGATATAAATTTTGACCTGTGAGAAC 2880
Db GAAACCTTTGGATCAGTATTGCTGCTCTAAAGTGATATAAATTTTGACCTGTGAGAAC 2880
Qy TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGTTTAAATCAACAGAAACACAGA 2940
Db TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGTTTAAATCAACAGAAACACAGA 2940
Qy ATTGATCATCATATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAAATGTGGTG 3000
Db ATTGATCATCATATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAAATGTGGTG 3000
Qy CCATGCGGGTCTACGGTTTGTGTAGTATCTAATACCTTAACCCACATGCTCAAAA 3060
Db CCATGCGGGTCTACGGTTTGTGTAGTATCTAATACCTTAACCCACATGCTCAAAA 3060
Qy TCAATGATACATATTTCTGAGAGCCCGCAATACCATTAAGATTTACTAAAAAANA 3120
Db TCAATGATACATATTTCTGAGAGCCCGCAATACCATTAAGATTTACTAAAAAANA 3120

RESULT 3

US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; FILE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12

;; PRIOR APPLICATION NUMBER: US 60/240,117
;; PRIOR FILING DATE: 2000-10-12
;; NUMBER OF SEQ ID NOS: 61
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 12
;; LENGTH: 4829
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match 99.6%; Score 3106.4; DB 3; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 1; Mismatches 1; Gaps 1;

Qy	1	AAGTGCTAAAGCCCTCCGAGCCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGCTAAAGCCCTCCGAGCCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTTCGCGCGCTGGGTTGTTCACCGCGCGCGCGCGCGAGGAAGCCACTGCAACCCAGGACCG	120
Db	61	CGTTTCGCGCGCTGGGTTGTTCACCGCGCGCGCGCGCGAGGAAGCCACTGCAACCCAGGACCG	120
Qy	121	GAGTGGAGCGCGCGAGCATGAAGCGCGAGCGCGCGCTTCATAGCGCACGTCGCGACGG	180
Db	121	GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCGCTTCATAGCGCACGTCGCGACGG	180
Qy	181	TCCGGCGCGCGCGGGGAGGAAATGCAATGCGCAGCAGCAATGGAACACAGAACAG	240
Db	181	TCCGGCGCGCGCGGGGAGGAAATGCAATGCGCAGCAGCAATGGAACACAGAACAG	240
Qy	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAGGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAGGATCGG	300
Qy	301	CCTAAATTGGAGCCCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGATGCTT	360
Db	301	CCTAAATTGGAGCCCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGATGCTT	360
Qy	361	GCCGATACAGAAATATCATGCTCATGATCGCTTAAGCCACCATGATTTTCAATGTT	420
Db	361	GCCGATACAGAAATATCATGCTCATGATCGCTTAAGCCACCATGATTTTCAATGTT	420
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTTATTCCTCCCATGCT	480
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTTATTCCTCCCATGCT	480
Qy	481	GGTGAGAACAGAGAAATACACTGTTTATTCGAAATTCGAAATTCGAAATTCGAAATTCGAA	540
Db	481	GGTGAGAACAGAGAAATACACTGTTTATTCGAAATTCGAAATTCGAAATTCGAAATTCGAA	540
Qy	541	GCAGTCTTAATGCTCTCTTGAAGCCCTCTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTGAAGCCCTCTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGAATGATATTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAAGTCGGA	660
Db	601	GGAATGATATTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAAGTCGGA	660
Qy	661	ATTGCTCTTACGATTATCACCAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA	720
Db	661	ATTGCTCTTACGATTATCACCAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA	720
Qy	721	ATTATACGTTAAAGATGGAGGCCACAGGATTTACGCAACAACTTTTAAGGCCCAAT	780
Db	721	ATTATACGTTAAAGATGGAGGCCACAGGATTTACGCAACAACTTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Db	841	GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900

Qy	901	GAAGAGAGACTCACTTATGTGCAATAGAGCTAGCCAAACATCGAAGAGATGCCAGATCA	960
Db	901	GAAGAGAGACTCACTTATGTGCAATAGAGCTAGCCAAACATCGAAGAGATGCCAGATCA	960
Qy	961	GCTGGAGTCCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Db	961	GCTGGAGTCCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Qy	1021	TGTCCAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATCTTATATGAGAA	1080
Db	1021	TGTCCAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATCTTATATGAGAA	1080
Qy	1081	AATGATGATCTGAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140
Db	1081	AATGATGATCTGAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140
Qy	1141	GCAGATTCATTTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	1200
Db	1141	GCAGATTCATTTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	1200
Qy	1201	TCAGAAATAATGATTTGATGCTGAAAGGAAGATCATAGATGTCATAGATAGGAATTAAT	1260
Db	1201	TCAGAAATAATGATTTGATGCTGAAAGGAAGATCATAGATGTCATAGATAGGAATTAAT	1260
Qy	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGATGCACTCT	1320
Db	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGATGCACTCT	1320
Qy	1321	GAGGAAATAATGCTTGGTCCCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTG	1380
Db	1321	GAGGAAATAATGCTTGGTCCCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTG	1380
Qy	1381	TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATCATGTTATGGAAGAGCAGAGACTC	1440
Db	1381	TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATCATGTTATGGAAGAGCAGAGACTC	1440
Qy	1441	ATTGAGTCACTGCTGATCTGTGACGCCACTAATTTATCTATGAGAAACAAACAGACATC	1500
Db	1441	ATTGAGTCACTGCTGATCTGTGACGCCACTAATTTATCTATGAGAAACAAACAGACATC	1500
Qy	1501	TGATAAATATCCATGACATCTTTTCATGTTTCCCAAAAGTCCAGAAAGAGGAAATTTGAG	1560
Db	1501	TGATAAATATCCATGACATCTTTTCATGTTTCCCAAAAGTCCAGAAAGAGGAAATTTGAG	1560
Qy	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATT	1620
Qy	1621	TTAAAGGAAAGCAAAATATAAAGCATCCAGTGGTGGGCTGCTCCCAAGTATTTTCAAG	1680
Db	1621	TTAAAGGAAAGCAAAATATAAAGCATCCAGTGGTGGGCTGCTCCCAAGTATTTTCAAG	1680
Qy	1681	TGTCCTCTCAAGAGAGAGATAGCAATTTACAGTGGTGGGCTGCTCCCAAGTATTTTCAAG	1740
Db	1681	TGTCCTCTCAAGAGAGAGATAGCAATTTACAGTGGTGGGCTGCTCCCAAGTATTTTCAAG	1740
Qy	1741	GGATCTAAATATCCAAAGTGGATGAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGAC	1800
Db	1741	GGATCTAAATATCCAAAGTGGATGAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGAC	1800
Qy	1801	TCCCTCTTATAGAGATCACTGTATCGTATGTTACGTTAAATCTGAGAGGTCACAGG	1860
Db	1801	TCCCTCTTATAGAGATCACTGTATCGTATGTTACGTTAAATCTGAGAGGTCACAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGCTAGTCACTGAGCTGTGACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGCTAGTCACTGAGCTGTGACTTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAAACAGAGAAATCCACATGTTGTTCCCTTTTAAAGCTATCAAGTCT	1980
Db	1921	AGTAAGTATAGTAAACAGAGAAATCCACATGTTGTTCCCTTTTAAAGCTATCAAGTCT	1980

QY 1981 GAAGATGACCACTTGCCAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACAGGT 2040
DB |||||
QY 1981 GAAGATGACCACTTGCCAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACAGGT 2040
DB |||||
QY 2041 CCTCTTCTGCTACTACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACA 2100
DB |||||
QY 2041 CCTCTTCTGCTACTACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACA 2100
DB |||||
QY 2101 TTGATGCGGATGCTCTACAGCTCATGATCTACAGCTGGGAAGAAATATCTCTACTGTG 2160
DB |||||
QY 2101 TTGATGCGGATGCTCTACAGCTCATGATCTACAGCTGGGAAGAAATATCTCTACTGTG 2160
DB |||||
QY 2161 CTGTTTCATATATGTGG--TCCTCAGGTGAGTGTGGTGAATATCGGTTTAAAGAGTCA 2218
DB |||||
QY 2161 CTGTTTCATATATGTGGTCTCTCAGGTGAGTGTGGTGAATATCGGTTTAAAGAGTCA 2220
DB |||||
QY 2219 AGTATTTCCGTTGAATACCTTACGCTCTCTAGGTTATGTGGTTGAGTATAGACAACA 2278
DB |||||
QY 2221 AGTATTTCCGTTGAATACCTTACGCTCTCTAGGTTATGTGGTTGAGTATAGACAACA 2280
DB |||||
QY 2279 GGGGATCTGTCAACGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAA 2338
DB |||||
QY 2281 GGGGATCTGTCAACGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAA 2340
DB |||||
QY 2339 TAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATG 2398
DB |||||
QY 2341 TAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATG 2400
DB |||||
QY 2399 ACTTAGATCGTGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAT 2458
DB |||||
QY 2401 ACTTAGATCGTGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAT 2460
DB |||||
QY 2459 TAATCCAGAGTTCAGATATCTTCAGGGTGTCTATTCCTGGGCCCGCTACTCTGTGGA 2518
DB |||||
QY 2461 TAATCCAGAGTTCAGATATCTTCAGGGTGTCTATTCCTGGGCCCGCTACTCTGTGGA 2520
DB |||||
QY 2519 TCTTCTATGATACAGGATACACGGAAGTGTATATGGGTCACTTACGACAGAAATGAACAGG 2578
DB |||||
QY 2521 TCTTCTATGATACAGGATACACGGAAGTGTATATGGGTCACTTACGACAGAAATGAACAGG 2580
DB |||||
QY 2579 GCTATTTACTTAGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTCTGAAACCAATCGTT 2638
DB |||||
QY 2581 GCTATTTACTTAGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTCTGAAACCAATCGTT 2640
DB |||||
QY 2639 TACTGCTCTTACATGGTTTCTCGATGGAATGTCCATTTTGACATACCAAGTATATTAC 2698
DB |||||
QY 2641 TACTGCTCTTACATGGTTTCTCGATGGAATGTCCATTTTGACATACCAAGTATATTAC 2700
DB |||||
QY 2699 TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGAC 2758
DB |||||
QY 2701 TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGAC 2760
DB |||||
QY 2759 ACAGCATAGAGTTTCTGAAATCGGAGAACATTTATGAATCTGATCTTTTGCACTACCTTC 2818
DB |||||
QY 2761 ACAGCATAGAGTTTCTGAAATCGGAGAACATTTATGAATCTGATCTTTTGCACTACCTTC 2820
DB |||||
QY 2819 AAGAAACCTTGGATCAGTATGCTGCTTAAAGTGATATAATTTTGACCTGTAGA 2878
DB |||||
QY 2821 AAGAAACCTTGGATCAGTATGCTGCTTAAAGTGATATAATTTTGACCTGTAGA 2880
DB |||||
QY 2879 ACTCTGCTGATACAGTGTATTTAAACCAATGAGGAGTTTATCAACAGAAACACA 2938
DB |||||
QY 2881 ACTCTGCTGATACAGTGTATTTAAACCAATGAGGAGTTTATCAACAGAAACACA 2940
DB |||||
QY 2939 GAATGATCATCATTTTGATACCTGCCATGTAACTCTCTGAAATTAATGTGG 2998
DB |||||
QY 2941 GAATGATCATCATTTTGATACCTGCCATGTAACTCTCTGAAATTAATGTGG 3000
DB |||||
QY 2999 TGCCATGAGGGGTCTACGGTTTGTGGTATGATTAATCTTAATACCTTAACCCCACTAGCTCAA 3058
DB |||||
QY 3001 TGCCATGAGGGGTCTACGGTTTGTGGTATGATTAATCTTAATACCTTAACCCCACTAGCTCAA 3060
DB |||||
QY 3059 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATAAGAAATTTACTAAAAA 3118
DB |||||

RESULT 4

US-10-982-512-12
; Sequence 12, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-12

Query Match 99.6%; Score 3106.4; DB 9; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 AAGTCTAAAGCTCCGAGGCCAAGCGGTGCTACTGCTCCGCGCTGCTTTAGTGCCG 60
DB 1 AAGTCTAAAGCTCCGAGGCCAAGCGGTGCTACTGCTCCGCGCTGCTTTAGTGCCG 60
QY 61 CGTTCGCGCTGGGTGTGTACCGCGCCCGCCGAGGAAGCACTGCAACACAGACCG 120
DB 61 CGTTCGCGCTGGGTGTGTACCGCGCCCGCCGAGGAAGCACTGCAACACAGACCG 120
QY 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCCGCTCCATAGCGCAGCTCGGACCG 180
DB 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCCGCTCCATAGCGCAGCTCGGACCG 180
QY 181 TCCGGCGCGCGCCGCGGGAAGGAAATGCAATGCGCAGCAATGGAACAGAACAG 240
DB 181 TCCGGCGCGCGCCGCGGGAAGGAAATGCAATGCGCAGCAATGGAACAGAACAG 240
QY 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTCGAATCAGAGGATCG 300
DB 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTCGAATCAGAGGATCG 300
QY 301 CTTAAATGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT 360
DB 301 CTTAAATGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT 360
QY 361 GCCGATACAGAAATATCATGGCTCATGATGGCTAAGGACCAACATGATTTTATGTTT 420
DB 361 GCCGATACAGAAATATCATGGCTCATGATGGCTAAGGACCAACATGATTTTATGTTT 420
QY 421 GTGAAGAGAAATGATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGCATGTCT 480
DB 421 GTGAAGAGAAATGATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGCATGTCT 480
QY 481 GGTGAGAAACAGAGAAATATACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA 540
DB 481 GGTGAGAAACAGAGAAATATACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA 540

Qy 491 GAGAAATACACGTGTTTATCTGAAATTCCTCAAAATCTATCAATAGACAGCAGTCTTAA 550
Db 506 GAGAAATACACGTGTTTATCTGAAATTCCTCAAAATCTATCAATAGACAGCAGTCTTAA 565
Qy 551 TGCTCTCTTGGAGCCCTCTTTGGATCTTTTTCAGGCAACACTCGACTATGGAATGTATT 610
Db 566 TGCTCTCTTGGAGCCCTCTTTGGATCTTTTTCAGGCAACACTCGACTATGGAATGTATT 625
Qy 611 CTCGAGAAGAAGAACTATTAAAGAGAAAGAAACGCAATTTGGAAACAGTCGGAAATTCCTTCT 670
Db 626 CTCGAGAAGAAGAACTATTAAAGAGAAAGAAACGCAATTTGGAAACAGTCGGAAATTCCTTCT 685
Qy 671 ACGATTATCAACGAAGAGTGGAAACATTTCTGTGTTTCAAGCCGGTAGTGGAAATTTATCACG 730
Db 686 ACGATTATCAACGAAGAGTGGAAACATTTCTGTGTTTCAAGCCGGTAGTGGAAATTTATCACG 745
Qy 731 TAAAGATGGAGGCCACAAAGGATTTACGCAACAACTTTTAAAGCCCAATCTAGTGGAAA 790
Db 746 TAAAGATGGAGGCCACAAAGGATTTACGCAACAACTTTTAAAGCCCAATCTAGTGGAAA 805
Qy 791 CTAGTTGTCTCCAAACATACGGATGGATCCAAATTTATGCCCGCTGTATCCAGACTGGATTG 850
Db 806 CTAGTTGTCTCCAAACATACGGATGGATCCAAATTTATGCCCGCTGTATCCAGACTGGATTG 865
Qy 851 CTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAAAGAGGAC 910
Db 866 CTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAAAGAGGAC 925
Qy 911 TCACATTATGGCAATAGCTAGCCAAACATGGAAAGAGATGCCAGATCAGCTGGAGTCG 970
Db 926 TCACATTATGGCAATAGCTAGCCAAACATGGAAAGAGATGCCAGATCAGCTGGAGTCG 985
Qy 971 CTACCTTTGTTCTCCAAAGAAATTTGTAGATATTCGGCTATTTGGTGCTGCCAAAG 1030
Db 986 CTACCTTTGTTCTCCAAAGAAATTTGTAGATATTCGGCTATTTGGTGCTGCCAAAG 1045
Qy 1031 CTGAAACAACTCCACGTGGTGTAAATTTCTTAAAGTTCTATATGAAGAAATATGATGAAT 1090
Db 1046 CTGAAACAACTCCACGTGGTGTAAATTTCTTAAAGTTCTATATGAAGAAATATGATGAAT 1105
Qy 1091 CTGAGGTGGAAATTTATGTTTACATCCCTATGTTGGAAACAAAGAGGCGAGATTCAT 1150
Db 1106 CTGAGGTGGAAATTTATGTTTACATCCCTATGTTGGAAACAAAGAGGCGAGATTCAT 1165
Qy 1151 TCCGTTATCTTAAACAGGTACAGCAATCTCTTAAAGTCACCTTTTAAAGATGTCCAGAAATA 1210
Db 1166 TCCGTTATCTTAAACAGGTACAGCAATCTCTTAAAGTCACCTTTTAAAGATGTCCAGAAATA 1225
Qy 1211 TGATTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACATAATTCACACCTTTTG 1270
Db 1226 TGATTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACATAATTCACACCTTTTG 1285
Qy 1271 AGATTCTATTTGAAGAGTTGAATATATTCAGAGCTGGATGGACTCTCTGAGGGGAAAT 1330
Db 1286 AGATTCTATTTGAAGAGTTGAATATATTCAGAGCTGGATGGACTCTCTGAGGGGAAAT 1345
Qy 1331 ATGCTTGGTCCATCTACTAGATCGCTCCAGCTCGCTACAGTAGTGTGATCTCAC 1390
Db 1346 ATGCTTGGTCCATCTACTAGATCGCTCCAGCTCGCTACAGTAGTGTGATCTCAC 1405
Qy 1391 CTGAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAAGCAGAGACTCATTTGAGTCAG 1450
Db 1406 CTGAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAAGCAGAGACTCATTTGAGTCAG 1465
Qy 1451 TGCTGTATCTGTGACGCCACTAAATTTATCTATGAAGAAACAAAGACATCTGGATAAATA 1510
Db 1466 TGCTGTATCTGTGACGCCACTAAATTTATCTATGAAGAAACAAAGACATCTGGATAAATA 1525
Qy 1511 TCCATGACATCTTTCATGTTTTTCCCAAGTCAGAGAGAAATTTGAGTTTATTTTTG 1570
Db 1526 TCCATGACATCTTTCATGTTTTTCCCAAGTCAGAGAGAAATTTGAGTTTATTTTTG 1585
Qy 1571 CCTCTGAATGCAAAACAGGTTTCCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAA 1630

Db 1586 CCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAA 1645
Qy 1631 GCAAAATATAAACGATCCAGTGGTGGCTCCCTGCTCCAAAGTGATTTCAAGTGTCTTATCA 1690
Db 1646 GCAAAATATAAACGATCCAGTGGTGGCTCCCTGCTCCAAAGTGATTTCAAGTGTCTTATCA 1705
Qy 1691 AAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTCTTTGGCGCGCATGGATCTAATA 1750
Db 1706 AAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTCTTTGGCGCGCATGGATCTAATA 1765
Qy 1751 TCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAG 1810
Db 1766 TCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAG 1825
Qy 1811 AGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGCAAGGCTGACTGACC 1870
Db 1826 AGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGCAAGGCTGACTGACC 1885
Qy 1871 GTGGCTACTCAATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTAAGTATATA 1930
Db 1886 GTGGCTACTCAATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTAAGTATATA 1945
Qy 1931 GTAAACAGAGAGATCCACACTGTGTCCCTTTTCAAGCTATCAAGTCTCTGGAAGATGACC 1990
Db 1946 GTAAACAGAGAGATCCACACTGTGTCCCTTTTCAAGCTATCAAGTCTCTGGAAGATGACC 2005
Qy 1991 CAATCTGCAAAAACAAAGGAAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCCTG 2050
Db 2006 CAATCTGCAAAAACAAAGGAAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCCTG 2065
Qy 2051 ACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTATGGGA 2110
Db 2066 ACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTATGGGA 2125
Qy 2111 TGCTCTACAAGCTCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGCTCATAT 2170
Db 2126 TGCTCTACAAGCTCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGCTCATAT 2185
Qy 2171 ATGCTGTCTCTCAGGTGCAAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230
Db 2186 ATGCTGTCTCTCAGGTGCAAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCT 2245
Qy 2231 TGAATACCTCCTAGCTCTCTAGGTATGTTGTTAGTATAGACAAACAGGGGATCTGTGTC 2290
Db 2246 TGAATACCTCCTAGCTCTCTAGGTATGTTGTTAGTATAGACAAACAGGGGATCTGTGTC 2305
Qy 2291 ACCGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAAAATAGAAATTCAGC 2350
Db 2306 ACCGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAAAATAGAAATTCAGC 2365
Qy 2351 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTG 2410
Db 2366 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTG 2425
Qy 2411 TGCGCATCCACGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGAGAGGT 2470
Db 2426 TGCGCATCCACGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGAGAGGT 2485
Qy 2471 CAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2530
Db 2486 CAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2545
Qy 2531 CAGGATACCGGAACGTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAG 2590
Db 2546 CAGGATACCGGAACGTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAG 2605
Qy 2591 GATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTCTGAAACCAAAATGCTTTACTGCTCTTAC 2650
Db 2606 GATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTCTGAAACCAAAATGCTTTACTGCTCTTAC 2665
Qy 2651 ATGCTTCTCTGGATGAGAAATGTCATTTTGGCAATACCAAGTATATTACTGAGTTTCTTTAG 2710

902	ACTTATGTGCA	CAATAGAGCT	AGCCACA	CTGAAGAAAGATG	CCAGATC	AGCTGGAGTCGCT	961
973	ACCTTTGTTT	CCTCAAGAAGAA	TTTGATAGAT	TATCTTGGCTA	TTTGCTGGTGTG	CCTCAAGAGCT	1033
962	ACCTTTGTTT	CCTCAGAAGAA	TTTGATAGAT	TATCTTGGCTA	TTTGCTGGTGTG	CCTCAAGAGCT	1021
1033	GAACAAC	TCTCCAGTGTG	TAAATTTCTT	TAGAAATTC	TATATGAAGAAAT	TGATGAATCT	1092
1022	GAACAAC	TCTCCAGTGTG	TAAATTTCTT	TAGAAATTC	TATATGAAGAAAT	TGATGAATCT	1081
1093	GAGGTGGAA	TTTATTCATG	TTTACATCCCT	TATGTTGGAAA	CAAGGAGGCGAGAT	TCATTCT	1152
1082	GAGGTGGAA	TTTATTCATG	TTTACATCCCT	TATGTTGGAAA	CAAGGAGGCGAGAT	TCATTCT	1141
1153	CGTTATCT	CTAAAA	CAGGTACAG	CAAAATCTTAA	AGTCACATTTTAA	GATGTCAGAAATAAATG	1212
1142	CGTTATCT	CTAAAA	CAGGTACAG	CAAAATCTTAA	AGTCACATTTTAA	GATGTCAGAAATAAATG	1201
1213	ATTGATGCT	GAAGGAGGAT	CA	TAGATGTCAT	A	TAGATAGGAACTAAATTCACCTTTTGAG	1272
1202	ATTGATGCT	GAAGGAGGAT	CA	TAGATGTCAT	A	TAGATAGGAACTAAATTCACCTTTTGAG	1261
1273	ATTCTATT	TTTGAAGGAG	TTGAA	TATATTG	CCAGAGCTGGAT	CGACTCCTCAGGGGAAAAATAT	1332
1262	ATTCTATT	TTTGAAGGAG	TTGAA	TATATTG	CCAGAGCTGGAT	CGACTCCTCAGGGGAAAAATAT	1321
1333	GCTTGGTCC	ATCTAGATCG	CTCCAGACT	CGCCTAC	AGATAGTTTGAT	CTCACTC	1392
1322	GCTTGGTCC	ATCTAGATCG	CTCCAGACT	CGCCTAC	AGATAGTTTGAT	CTCACTC	1381
1393	GAATATT	TTTATCC	CAGTGAAG	ATGATGTTAT	CGAAAAGGCGAGAG	ACTCAITTAGTCAAGTG	1452
1382	GAATATT	TTTATCC	CAGTGAAG	ATGATGTTAT	CGAAAAGGCGAGAG	ACTCAITTAGTCAAGTG	1441
1453	CCTGATTCT	GTGACGCC	ACTAA	TTATCTAT	GAAGAAAC	CAACAGACATCTCGATAAAATATC	1512
1442	CCTGATTCT	GTGACGCC	ACTAA	TTATCTAT	GAAGAAAC	CAACAGACATCTCGATAAAATATC	1501
1513	CATGACAT	CTTTTCATG	TTTTTCCC	CAAGTCCAG	AAAGAGAAATTTAG	TTTTTTTGCC	1572
1502	CATGACAT	CTTTTCATG	TTTTTCCC	CAAGTCCAG	AAAGAGAAATTTAG	TTTTTTTGCC	1561
1573	TCTGAATG	CAAAACAG	GTTC	CCGTCATTT	TATACAAATTA	CAATCTATTTTAAAGGAAAGC	1632
1562	TCTGAATG	CAAAACAG	GTTC	CCGTCATTT	TATACAAATTA	CAATCTATTTTAAAGGAAAGC	1621
1633	AAATATAAC	GCATCCAG	TGTGG	CTGCTCC	AAAGTGATTTCA	AGTGTCCTATCAAA	1692
1622	AAATATAAC	GCATCCAG	TGTGG	CTGCTCC	AAAGTGATTTCA	AGTGTCCTATCAAA	1681
1693	GAGGAGAT	GACAA	TTTACC	AGTGGTGAAT	TGGAAAGTTCTT	GGCGGCGATGGAATCTAAATATC	1752
1682	GAGGAGAT	GACAA	TTTACC	AGTGGTGAAT	TGGAAAGTTCTT	GGCGGCGATGGAATCTAAATATC	1741
1753	CAAGTTGAT	GAAATCAG	AAGGCTGG	TATATTTT	GAAGGCA	CCAAAGACTCCCTCTTTAGAG	1812
1742	CAAGTTGAT	GAAATCAG	AAGGCTGG	TATATTTT	GAAGGCA	CCAAAGACTCCCTCTTTAGAG	1801
1813	CATCACC	TGTAGCTAG	TTACG	TATAGTGG	AGATGTC	CAAGGCTGAC	1872
1802	CATCACC	TGTAGCTAG	TTACG	TATAGTGG	AGATGTC	CAAGGCTGAC	1861
1873	GGCTACT	CTCACAT	TTTGTG	CTCAGT	CAGCACTGTG	ACTCTTTTATAAGTAAGTATAGT	1932
1862	GGCTACT	CTCACAT	TTTGTG	CTCAGT	CAGCACTGTG	ACTCTTTTATAAGTAAGTATAGT	1921
1933	AACCA	GAGAAATCC	ACTGTGTG	CTTTC	CAAGCTATCA	AGTCTTCAAGTCAAGGATGACCCA	1992
1922	AACCA	GAGAAATCC	ACTGTGTG	CTTTC	CAAGCTATCA	AGTCTTCAAGTCAAGGATGACCCA	1981
1993	ACTTGC	AAAA	CAAAAGAA	TTTTTGG	CCACCA	TTTTTGGATTCAGCAGGCTCTCTTCCTGAC	2052
1982	ACTTGC	AAAA	CAAAAGAA	TTTTTGG	CCACCA	TTTTTGGATTCAGCAGGCTCTCTTCCTGAC	2041

Qy	2053	TATACTCCTCAGAAAATTTTCTCTTTTGAAGTACTACTCGGATTTACAAATTTACAAATCTGATGGGATG	2111
Db	2042	TATACTCCTCAGAAAATTTTCTCTTTTGAAGTACTACTCGGATTTACAAATTTACAAATCTGATGGGATG	2101
Qy	2113	CTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTTACTGTGCTGCTTCATATAT	2172
Db	2102	CTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTTACTGTGCTGCTTCATATAT	2161
Qy	2173	GGTGGTCTCAGGTGCAGTTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCGGCTTG	2232
Db	2162	GGTGGTCTCAGGTGCAGTTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCGGCTTG	2221
Qy	2233	AATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCCTGTCCAC	2292
Db	2222	AATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCCTGTCCAC	2281
Qy	2293	CGAGGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTAAGCAT	2352
Db	2282	CGAGGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTAAGCAT	2341
Qy	2353	CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTG	2412
Db	2342	CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTG	2401
Qy	2413	GGCATCCAGCGCTGGTCTTATGAGAGATACCTCTCCCTGATCGCATTAATCGAGAGTCA	2472
Db	2402	GGCATCCAGCGCTGGTCTTATGAGAGATACCTCTCCCTGATCGCATTAATCGAGAGTCA	2461
Qy	2473	GATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACA	2532
Db	2462	GATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACA	2521
Qy	2533	GGATACACGGAACGTTATATGGGTCACTTCAGCAGAGATGAACAGGGCTTATTACTTAGGA	2592
Db	2522	GGATACACGGAACGTTATATGGGTCACTTCAGCAGAGATGAACAGGGCTTATTACTTAGGA	2581
Qy	2593	TCGTGGCCATGCAAGCAGAAAAGTCCCTCTCTGAAACCAATCGTTTACTGCTCTTACAT	2652
Db	2582	TCGTGGCCATGCAAGCAGAAAAGTCCCTCTCTGAAACCAATCGTTTACTGCTCTTACAT	2641
Qy	2653	GGTTTCTCGGATGAGAAATGCTCCATTTTGGCATACACAGTATATTACTGAGTTTTTATAGT	2712
Db	2642	GGTTTCTCGGATGAGAAATGCTCCATTTTGGCATACACAGTATATTACTGAGTTTTTATAGT	2701
Qy	2713	AGGGCTGGAAGGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGATT	2772
Db	2702	AGGGCTGGAAGGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGATT	2761
Qy	2773	CCTGAATCGGGAGAACATTATGAACCTGATCTTTTGCACCTACCTTCAAGAAACCTTGGA	2832
Db	2762	CCTGAATCGGGAGAACATTATGAACCTGATCTTTTGCACCTACCTTCAAGAAACCTTGGA	2821
Qy	2833	TCACGTTATGCTGCTCTTAAAGTCAATATTTTGGACCTGTGTAGNACTCTCTGGTATAC	2892
Db	2822	TCACGTTATGCTGCTCTTAAAGTCAATATTTTGGACCTGTGTAGNACTCTCTGGTATAC	2881
Qy	2893	ACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTCATCATCAC	2952
Db	2882	ACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTCATCATCAC	2941
Qy	2953	ATTTTGTATACCTGCCATGTAAACATCTACTCCTGAAAATAAATGTGTGCTGCAATGAGGGT	3012
Db	2942	ATTTTGTATACCTGCCATGTAAACATCTACTCCTGAAAATAAATGTGTGCTGCAATGAGGGT	3001
Qy	3013	CTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACA	3072
Db	3002	CTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACA	3061
Qy	3073	TATTTCTGAGAGACCCAGCAATACCAATAGAAATTTCTTAAAAAAA	3117
Db	3062	TATTTCTGAGAGACCCAGCAATACCAATAGAAATTTCTTAAAAAAA	3106

RESULT 7		US-09-976-674-22	
		; Sequence 22, Application US/09976674	
		; Patent No. US20020115843A1	
		; GENERAL INFORMATION:	
		; APPLICANT: Qi, Steve	
		; APPLICANT: Akinsanya, Karen	
		; APPLICANT: Riviere, Pierre	
		; APPLICANT: Junien, Jean-Louis	
		; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV	
		; FILE REFERENCE: 70669	
		; CURRENT APPLICATION NUMBER: US/09/976,674	
		; PRIOR FILING DATE: 2001-10-12	
		; PRIOR APPLICATION NUMBER: US 60/240,117	
		; PRIOR FILING DATE: 2000-10-12	
		; NUMBER OF SEQ ID NOS: 61	
		; SOFTWARE: PatentIn version 3.1	
		; SEQ ID NO 22	
		; LENGTH: 4685	
		; TYPE: DNA	
		; ORGANISM: Homo sapiens	
		US-09-976-674-22	
		Query Match	
		Best Local Similarity 90.5%; Score 2824.4; DB 3; Length 4685;	
		Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;	
Qy	1	AAGTCTAAGCCTCCGAGGCAAGGCGCTGCTACTGCGCGGCTGCTTCTTAGTGCGG	60
Db	1	AAGTCTAAGCCTCCGAGGCAAGGCGCTGCTACTGCGCGGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTGCGCGCTGGGTTGTCAACGCGCGCCGCCGAGGAAGGCACTGGCAACGAGACG	120
Db	61	CGTTGCGCGCTGGGTTGTCAACGCGCGCCGCCGAGGAAGGCACTGGCAACGAGACG	120
Qy	121	GAGTGGAGCGCGCAGCATGAAGCGCGCGAGCGCCCGCTCCATAGCGCACTGTCGGGACG	180
Db	121	GAGTGGAGCGCGCAGCATGAAGCGCGCGAGCGCCCGCTCCATAGCGCACTGTCGGGACG	180
Qy	181	TCGGGCGCGCGCGGGGGAAGGAAATGCAATCGCAGCGCAATGGAAACAGAACAG	240
Db	181	TCGGGCGCGCGCGGGGGAAGGAAATGCAATCGCAGCGCAATGGAAACAGAACAG	240
Qy	241	CTGGGTGTGAGATATTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAGAGGATCGG	300
Db	241	CTGGGTGTGAGATATTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAGAGGATCGG	300
Qy	301	CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT	360
Db	301	CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT	360
Qy	361	GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCAACCATGATTTTCATGTTT	420
Db	361	GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCAACCATGATTTTCATGTTT	420
Qy	421	GTGAGAGGAATGATCCAGATGGACCTTATTCAGACAGATCTATTACCTTGCCATGCT	480
Db	421	GTGAGAGGAATGATCCAGATGGACCTTATTCAGACAGATCTATTACCTTGCCATGCT	480
Qy	481	GGTGAGACAGAAATATACCTGTTTATTCGAAATCCGAAACTATCAATAGAGCA	540
Db	481	GGTGAGACAGAAATATACCTGTTTATTCGAAATCCGAAACTATCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGCACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGCACTAT	600
Qy	601	GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATTCGAAACAGTCGGA	660
Db	601	GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATTCGAAACAGTCGGA	660
Qy	661	ATTGCTTCTTACGATTATCAACAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA	720

Db	661	ATTGCTTCTTACGATTATCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA	720
Qy	721	ATTATACCTTAAAGATGGAGGCAACAGGATTTTACGCAACAGCTTTTAAAGCCCAAT	780
Db	721	ATTATACCTTAAAGATGGAGGCAACAGGATTTTACGCAACAGCTTTTAAAGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTCCTCAACATACGGAATGATCCAAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTTGTCCTCAACATACGGAATGATCCAAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGGATTGCTTTTATACATAGCAACGATATTGGATATCTAACATCGTAAACAGAGAA	900
Db	841	GACTGGATTGCTTTTATACATAGCAACGATATTGGATATCTAACATCGTAAACAGAGAA	900
Qy	901	GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960
Db	901	GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960
Qy	961	GCTGGAGTCGCTACCTTTGTTTCTCCAAAGAAAGATTTTGATAGATATTCTGGCTATTGGTGG	1020
Db	961	GCTGGAGTCGCTACCTTTGTTTCTCCAAAGAAAGATTTTGATAGATATTCTGGCTATTGGTGG	1020
Qy	1021	TGTCCTAAAGGCTGAAACAACTCCCAAGTGGTAAATTTCTTAGAATTTCTATATGAAGAA	1080
Db	1021	TGTCCTAAAGGCTGAAACAACTCCCAAGTGGTAAATTTCTTAGAATTTCTATATGAAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGGAATTTTTCATGTTTACATCCCTTATGTTGGAACAAGGAGG	1140
Db	1081	AATGATGAATCTGAGGTGGAATTTTTCATGTTTACATCCCTTATGTTGGAACAAGGAGG	1140
Qy	1141	GCAGATTCATTCGTTATCTTAAACAGGTACAGCAATCTTAAGTCACCTTTTAAAGATG	1200
Db	1141	GCAGATTCATTCGTTATCTTAAACAGGTACAGCAATCTTAAGTCACCTTTTAAAGATG	1200
Qy	1201	TCAGAAATATGATTGATGCTGAGGAGGATCATAGATGCTATAGTAAGGAACCTAAT	1260
Db	1201	TCAGAAATATGATTGATGCTGAGGAGGATCATAGATGCTATAGTAAGGAACCTAAT	1260
Qy	1261	CAACCTTTGAGATCTTATTTGAAAGGAGTTGAATATATGCCAGAGCTGGATGACTCCT	1320
Db	1261	CAACCTTTGAGATCTTATTTGAAAGGAGTTGAATATATGCCAGAGCTGGATGACTCCT	1320
Qy	1321	GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG	1380
Db	1321	GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAAATTTATCCAGTAGAAGATGATGTTATGGAAGGCAAGAGACTC	1440
Db	1381	TTGATCTCACCTGAAATTTATCCAGTAGAAGATGATGTTATGGAAGGCAAGAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATTTCTGTGACCGCACTAATTTATCTATGAAGAAACAAACAGATC	1500
Db	1441	ATTGAGTCAGTGCCTGATTTCTGTGACCGCACTAATTTATCTATGAAGAAACAAACAGATC	1500
Qy	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCCAGAAAGGAAATTTGAG	1560
Db	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCCAGAAAGGAAATTTGAG	1560
Qy	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTTCGTCATTTATACAAAATTTACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTTCGTCATTTATACAAAATTTACATCTATT	1620
Qy	1621	TTAAAGGAAGCAATATAACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG	1680
Db	1621	TTAAAGGAAGCAATATAACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG	1680
Qy	1681	TGTCCTATCAAGAGGATAGCAATTTACAGTGGTGAATGGGAGGTTCTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAGAGGATAGCAATTTACAGTGGTGAATGGGAGGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCCAAAGTGTGATGAAGTCCAGAGGCTGGTATATTTTTGAAGGCCAACAGAC	1800

Db 1741 GGATCTAATATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATATTTTGAAGCCACCAAGAC 1800
Qy 1801 TCCCTCTTTAGAGCATCACCTCTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTCTTTAGAGCATCACCTCTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCAATCTTCTGCTGCAATCAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCAATCTTCTGCTGCAATCAGTCAGCACTGTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCT 1980
Db 1921 AGTAAGTATAGTAAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCT 1980
Qy 1981 GAAAGATGACCCAACTTGCACAAACAAAGGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGT 2040
Db 1981 GAAAGATGACCCAACTTGCACAAACAAAGGAAATTTTGGGCCACCACTTTTGGATTTC 2034
Qy 2041 CCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACA 2100
Db 2035 ----- 2034
Qy 2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG 2160
Db 2035 ----- 2034
Qy 2161 CTGTTTCATATATGGTGGTCTCAGTGCAGTGTGTGTAATAATCGGTTTAAAGAGTCAAG 2220
Db 2035 -----GTCCCTCAGGTGCAGTGTGTGTAATAATCGGTTTAAAGAGTCAAG 2078
Qy 2221 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGTGTAGTGATAGACAACAGG 2280
Db 2079 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGTGTAGTGATAGACAACAGG 2138
Qy 2281 GGATCTGTCCAGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGTCAAATA 2340
Db 2139 GGATCTGTCCAGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA 2198
Qy 2341 GAAATTTGACATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2199 GAAATTTGACATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC 2258
Qy 2401 TTAGATCGTGTGGGCATCCAGGCTGTCTTATGAGGATACCTCTCCCTGTATGGCAATTA 2460
Db 2259 TTAGATCGTGTGGGCATCCAGGCTGTCTTATGAGGATACCTCTCCCTGTATGGCAATTA 2318
Qy 2461 ATGCAGAGGTCAAGATATCTTCAGGTTGCTATTCGTGGGCCCCAGTCACTCTGTGGATC 2520
Db 2319 ATGCAGAGGTCAAGATATCTTCAGGTTGCTATTCGTGGGCCCCAGTCACTCTGTGGATC 2378
Qy 2521 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACCCAGAGATGAACAGGGC 2580
Db 2379 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACCCAGAGATGAACAGGGC 2438
Qy 2581 TATTACTTATAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTA 2640
Db 2439 TATTACTTATAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTA 2498
Qy 2641 CTGCTCTTACATGGTTTCTCGGATGAGAAATGCTCAATTTTGCACATACCACTATTAATCTG 2700
Db 2499 CTGCTCTTACATGGTTTCTCGGATGAGAAATGCTCAATTTTGCACATACCACTATTAATCTG 2558
Qy 2701 AGTTTATTTAGTGAAGGCTGGAAGCCATATGATTTACATGATCTATCTCAGGAGAGACAC 2760
Db 2559 AGTTTATTTAGTGAAGGCTGGAAGCCATATGATTTACATGATCTATCTCAGGAGAGACAC 2618
Qy 2761 AGCATTAAGTCTTCTGAATCGGAGAACATTATGATGATCTTTTGGCACTACCTTTCAA 2820
Db 2619 AGCATTAAGTCTTCTGAATCGGAGAACATTATGATGATCTTTTGGCACTACCTTTCAA 2678
Qy 2821 GAAACCTTTGGATCAGCTATTTGCTCTCTTAAAGTGATATAAATTTTGACCTGTGTAGAAC 2880
Db 2679 GAAACCTTTGGATCAGCTATTTGCTCTCTTAAAGTGATATAAATTTTGACCTGTGTAGAAC 2738

RESULT 8

US-10-982-512-22
; Sequence 22, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-22

Query Match 90.5%; Score 2824.4; DB 9; Length 4685;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;

Qy 1 AAGTGCTAAAGCCCTCCGAGGCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTCTAGTCCG 60
Db 1 AAGTGCTAAAGCCCTCCGAGGCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTCTAGTCCG 60
Qy 61 CGTTCCGCGCTGGGTTGTACCGCGCGCGCGCGAGAGCCACTGCAACACGAGGACCG 120
Db 61 CGTTCCGCGCTGGGTTGTACCGCGCGCGCGCGAGAGCCACTGCAACACGAGGACCG 120
Qy 121 GAGTGAGGCGCGCGAGCATGAAGCGCGCGCGCTCCATAGCGCACGTCGCGGACG 180
Db 121 GAGTGAGGCGCGCGAGCATGAAGCGCGCGCGCTCCATAGCGCACGTCGCGGACG 180
Qy 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCGCAATGGAAACAGAA 240
Db 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCGCAATGGAAACAGAA 240
Qy 241 CTGGGTTTGCAGATATTTGAACTGCGGACTGTGAGGAGATATTGAATCAGGATCGG 300
Db 241 CTGGGTTTGCAGATATTTGAACTGCGGACTGTGAGGAGATATTGAATCAGGATCGG 300
Qy 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTTAAAAGCTGCTT 360
Db 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTTAAAAGCTGCTT 360
Qy 361 GCCGATACCAAGAAATATCATGATGGCTATCATGAGGTAAGGCAACCATGATTTTCAATGTT 420
Db 361 GCCGATACCAAGAAATATCATGATGGCTATCATGAGGTAAGGCAACCATGATTTTCAATGTT 420

Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACAAATCGTTTA 2640
Db |||||
Qy 2439 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACAAATCGTTTA 2498
Db |||||
Qy 2641 CTGCTCTTACATGGTTTCTCGATGAGATGTCCATTTTGGACATACCATATATTACTG 2700
Db |||||
Qy 2499 CTGCTCTTACATGGTTTCTCGATGAGATGTCCATTTTGGACATACCATATATTACTG 2558
Db |||||
Qy 2701 AGTTTTTTAGTAGGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 2760
Db |||||
Qy 2559 AGTTTTTTAGTAGGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 2618
Db |||||
Qy 2761 AGCATAGAGTTTCTCGAATCGGGAGAACATTATGAACTGCAATCTTTTGGACTACCTTCAA 2820
Db |||||
Qy 2619 AGCATAGAGTTTCTCGAATCGGGAGAACATTATGAACTGCAATCTTTTGGACTACCTTCAA 2678
Db |||||
Qy 2821 GAAAACTTGGATCAGTATTCGTCTCTAAAGTGATATAATTTTGGACCTGTGTAGAAC 2880
Db |||||
Qy 2679 GAAAACTTGGATCAGTATTCGTCTCTAAAGTGATATAATTTTGGACCTGTGTAGAAC 2738
Db |||||
Qy 2881 TCTCTGATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA 2940
Db |||||
Qy 2739 TCTCTGATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA 2798
Db |||||
Qy 2941 ATTGATCATCACATTTTGTATCTGCTGCAATGTAACTCTCTCTGAAAAATAATGTGGTG 3000
Db |||||
Qy 2799 ATTGATCATCACATTTTGTATCTGCTGCAATGTAACTCTCTCTGAAAAATAATGTGGTG 2858
Db |||||
Qy 3001 CCATGACGGGCTTACGGTTTGTGTAGTAACTCTAAATACCTTAAACCCACATGCTCAAAA 3060
Db |||||
Qy 2859 CCATGACGGGCTTACGGTTTGTGTAGTAACTCTAAATACCTTAAACCCACATGCTCAAAA 2918
Db |||||
Qy 3061 TCAATGATACATATTCCTGAGAGACCCAGCAATACCAATAGATTACTTAAAAAATAAAAA 3120
Db |||||
Qy 2919 TCAATGATACATATTCCTGAGAGACCCAGCAATACCAATAGATTACTTAAAAAATAAAAA 2978
Db |||||

RESULT 9

US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20

Query Match 89.98; Score 2806.4; DB 3; Length 4676;
Best Local Similarity 95.18; Pred. No. 0;
Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;
Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGTCTTCTTAGTGCCG 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGTCTTCTTAGTGCCG 60
Qy 61 CGTTGCGCGCGCTGGTTGTTCACCGCGCGCGCGCGAGAGCCACTGCAACCAAGGACCG 120
Db 61 CGTTGCGCGCGCTGGTTGTTCACCGCGCGCGCGCGAGAGCCACTGCAACCAAGGACCG 120
Qy 121 GAGTGGAGGGCGGCGACGATGAAGAGGGCGGCGCGCTCCATAGCGCACGTCGGGACGG 180

Db 121 GAGTGGAGGGCGGCGACGATGAAGAGGGCGGCGCGCTCCATAGCGCACGTCGGGACGG 180
Qy 181 TCCGGGCGGGCGGGGGAGGAAATTCACCAATGCGCAGCAGCAATGGAAAAACAGAACAG 240
Db |||||
Qy 181 TCCGGGCGGGCGGGGGAGGAAATTCACCAATGCGCAGCAGCAATGGAAAAACAGAACAG 240
Db |||||
Qy 241 CTGGGTGTTGAGATATTTGAAACTCGGGACTGTGAGGAGATATTTGAATCACAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTCGGGACTGTGAGGAGATATTTGAATCACAGGATCGG 300
Qy 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGAGCTCAGCTTAAAGAGCTGCTT 360
Db 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGAGCTCAGCTTAAAGAGCTGCTT 360
Qy 361 GCCGATACACAGAAAAATATCATGGCTACATGATGGCTAAGGCCACCAATGATTTTCAATGTTT 420
Db 361 GCCGATACACAGAAAAATATCATGGCTACATGATGGCTAAGGCCACCAATGATTTTCAATGTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCATTTTCAGACAGAACTCTATTACCTTGGCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTTTCAGACAGAACTCTATTACCTTGGCATGCT 480
Qy 481 GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Qy 541 GCAGTCTTAAATGCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCCAACACTGAGCTAT 600
Db 541 GCAGTCTTAAATGCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCCAACACTGAGCTAT 600
Qy 601 GGAATGTATTTCTCGAGAGAGAACTATTAAGAGAGAAAGAAACGCGATTGGAAACAGTCGGA 660
Db 601 GGAATGTATTTCTCGAGAGAGAACTATTAAGAGAGAAAGAAACGCGATTGGAAACAGTCGGA 660
Qy 661 ATTGCTTTTACGATTTATCACCAAGAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGGA 720
Db 661 ATTGCTTTTACGATTTATCACCAAGAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGGA 720
Qy 721 ATTTATCACGCTAAAGATGAGGGCCACAAAGGATTTTACGCAACCACTTTTAAAGGCCCAAT 780
Db 721 ATTTATCACGCTAAAGATGAGGGCCACAAAGGATTTTACGCAACCACTTTTAAAGGCCCAAT 780
Qy 781 CTAGTGGAAACTAGTTGTCCCAACATACGATGGATGCCAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAAACTAGTTGTCCCAACATACGATGGATGCCAAATTTATGCCCGCTGATCCA 840
Qy 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTTAACTGTAACCCAGAGAA 900
Db 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTTAACTGTAACCCAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTCACATGAGCTAGCCCAATGGAAGAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTCACATGAGCTAGCCCAATGGAAGAGAGATGCCAGATCA 960
Qy 961 GCTGAGTCTGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGAGTCTGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGTGGGAAATTTATCATGTTTACATCCCTATCTGCGAAACAGGAGG 1140
Db 1081 AATGATGAATCTGAGTGGGAAATTTATCATGTTTACATCCCTATCTGTTGAAACAGGAGG 1140
Qy 1141 GCAGATTCAATTCGCTTATCTTAAACAGAGTACAGCAAACTCTTAAAGTCACTTTTAAAGTG 1200
Db 1141 GCAGATTCAATTCGCTTATCTTAAACAGAGTACAGCAAACTCTTAAAGTCACTTTTAAAGTG 1200
Qy 1201 TCAGAAATTAATGATTTGATGCTGAAGGAGGATCATAGATGTGCATAGTAAGGAACTAAT 1260
Db |||||

Db	1201	TCAGAAATPAATGATGTGCTGAAGGAGGATCATATAGATGTCTATAGATAAGAACTAAAT	1260
Qy	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTGAAATATATTTGCGAGAGCTGGATGAGCTCCT	1320
Db	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTGAAATATATTTGCGAGAGCTGGATGAGCTCCT	1320
Qy	1321	GAGGAAATATATGCTTGGTCCATCCTACTAGATCGCTCCAGACCTCGCCTACAGATAGTG	1380
Db	1321	GAGGAAATATATGCTTGGTCCATCCTACTAGATCGCTCCAGACCTCGCCTACAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGATC	1440
Db	1381	TTGATCTCACCTGAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGATC	1440
Qy	1441	ATTGATCAGTGCCTGATCTGTGAGCGCCACTAAATATCTATGAAGAAACACAGACATC	1500
Db	1441	ATTGATCAGTGCCTGATCTGTGAGCGCCACTAAATATCTATGAAGAAACACAGACATC	1500
Qy	1501	TGGATAAATATCCATGACATCTTTCATGTTTCCCAAGTCACGAGAGGAAATTCAG	1560
Db	1501	TGGATAAATATCCATGACATCTTTCATGTTTCCCAAGTCACGAGAGGAAATTCAG	1560
Qy	1561	TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCATTTATACAAATTTACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCATTTATACAAATTTACATCTATT	1620
Qy	1621	TTAAAGGAAGCAATATAAAGATGAAAGTCCAGTGGTGGCTGCTCTCCAGTGATTTCAAG	1680
Db	1621	TTAAAGGAAGCAATATAAAGATGAAAGTCCAGTGGTGGCTGCTCTCCAGTGATTTCAAG	1680
Qy	1681	TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGAAGTTCTTGGCGGGAT	1740
Db	1681	TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGAAGTTCTTGGCGGGAT	1740
Qy	1741	GGATCTAATATCAAGTTGATGAAAGTCCAGTGGTGGTATATTTTGAAGGCCACCAAGAC	1800
Db	1741	GGATCTAATATCAAGTTGATGAAAGTCCAGTGGTGGTATATTTTGAAGGCCACCAAGAC	1800
Qy	1801	TCCCTTTTAGAGCATCACTCTAGTACATTTTGTGTCATCAGTCAGCACTGTGACTTTTATA	1860
Db	1801	TCCCTTTTAGAGCATCACTCTAGTACATTTTGTGTCATCAGTCAGCACTGTGACTTTTATA	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTTGTGTCATCAGTCAGCACTGTGACTTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTTGTGTCATCAGTCAGCACTGTGACTTTTATA	1920
Qy	1921	AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCT	1980
Db	1921	AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCT	1980
Qy	1981	GAAGATGACCCAACTTGCMAAACAAAGAAATTTTGGCCACCAATTTTGGATTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTTGCMAAACAAAGAAATTTTGGCCACCAATTTTGGATTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACA	2100
Db	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACA	2100
Qy	2101	TTGTATGGATGCTCTTACAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG	2160
Db	2101	TTGTATGGATGCTCTTACAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTGGTCTCCTCAGGTGAGTGTGTAATATCGTTTAAAGAGTCAAG	2220
Db	2161	CTGTTTCATATATGTGGTCTCCTCAGGTGAGTGTGTAATATCGTTTAAAGAGTCAAG	2220
Qy	2221	TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGATGATAGACACAGG	2280
Db	2281	GGATCCTGTCAACCGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAATA	2340
Db	2180	-----GGGTCAAATA	2180
Qy	2341	GAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTAATTCATTGAC	2400
Db	2341	GAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTAATTCATTGAC	2400
Qy	2390	GAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTAATTCATTGAC	2450
Db	2390	GAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTAATTCATTGAC	2450
Qy	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGCAATTA	2460
Db	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGCAATTA	2460
Qy	2450	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGCAATTA	2510
Db	2450	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGCAATTA	2510
Qy	2461	ATGCAAGGCTCAGATATCTTTCAGGTTGCTTATGCTGGGGCCCGGCTGCTGCTGATC	2520
Db	2461	ATGCAAGGCTCAGATATCTTTCAGGTTGCTTATGCTGGGGCCCGGCTGCTGCTGATC	2520
Qy	2510	ATGCAAGGCTCAGATATCTTTCAGGTTGCTTATGCTGGGGCCCGGCTGCTGCTGATC	2570
Db	2510	ATGCAAGGCTCAGATATCTTTCAGGTTGCTTATGCTGGGGCCCGGCTGCTGCTGATC	2570
Qy	2521	TTCTATGATACAGGATACACGGAAACGTTATATGGGTTCACCTGACACAGAAATGAAACGGC	2580
Db	2521	TTCTATGATACAGGATACACGGAAACGTTATATGGGTTCACCTGACACAGAAATGAAACGGC	2580
Qy	2570	TTCTATGATACAGGATACACGGAAACGTTATATGGGTTCACCTGACACAGAAATGAAACGGC	2630
Db	2570	TTCTATGATACAGGATACACGGAAACGTTATATGGGTTCACCTGACACAGAAATGAAACGGC	2630
Qy	2581	TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTA	2640
Db	2581	TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTA	2640
Qy	2630	TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTA	2690
Db	2630	TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTA	2690
Qy	2641	CTGCTCTTACATGCTTCTGCTGGATGGAATGTCCATTTTGCACATACCAAGTATTTACTG	2700
Db	2641	CTGCTCTTACATGCTTCTGCTGGATGGAATGTCCATTTTGCACATACCAAGTATTTACTG	2700
Qy	2690	CTGCTCTTACATGCTTCTGCTGGATGGAATGTCCATTTTGCACATACCAAGTATTTACTG	2750
Db	2690	CTGCTCTTACATGCTTCTGCTGGATGGAATGTCCATTTTGCACATACCAAGTATTTACTG	2750
Qy	2701	AGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2760
Db	2701	AGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2760
Qy	2750	AGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2810
Db	2750	AGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2810
Qy	2761	AGCATAAGAGTTCTCTGAATCGGAGAAACATTTATGAACTGTCATCTTTTGCACTACCTTCAA	2820
Db	2761	AGCATAAGAGTTCTCTGAATCGGAGAAACATTTATGAACTGTCATCTTTTGCACTACCTTCAA	2820
Qy	2810	AGCATAAGAGTTCTCTGAATCGGAGAAACATTTATGAACTGTCATCTTTTGCACTACCTTCAA	2870
Db	2810	AGCATAAGAGTTCTCTGAATCGGAGAAACATTTATGAACTGTCATCTTTTGCACTACCTTCAA	2870
Qy	2821	GAAACCTTGGATCAGCTATTTGCTCTTAAAGTGTATATAATTTTGAAGTGTGTAAGAC	2880
Db	2821	GAAACCTTGGATCAGCTATTTGCTCTTAAAGTGTATATAATTTTGAAGTGTGTAAGAC	2880
Qy	2870	GAAACCTTGGATCAGCTATTTGCTCTTAAAGTGTATATAATTTTGAAGTGTGTAAGAC	2940
Db	2870	GAAACCTTGGATCAGCTATTTGCTCTTAAAGTGTATATAATTTTGAAGTGTGTAAGAC	2940
Qy	2881	TCTCTGTTATACATCTGCTTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA	2940
Db	2881	TCTCTGTTATACATCTGCTTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA	2940
Qy	2940	TCTCTGTTATACATCTGCTTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA	3000
Db	2940	TCTCTGTTATACATCTGCTTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA	3000
Qy	2990	CCATGCAAGGCTCTACGTTTGTGGTAGTAAATCTTAATACCTTAAACCCCATGCTCAAAA	3060
Db	2990	CCATGCAAGGCTCTACGTTTGTGGTAGTAAATCTTAATACCTTAAACCCCATGCTCAAAA	3060
Qy	3061	TCAATATGATACATATCTCTGAGAGACCCAGCAATACCAATAGAAATTAATAAAAAAAA	3120
Db	3061	TCAATATGATACATATCTCTGAGAGACCCAGCAATACCAATAGAAATTAATAAAAAAAA	3120
Qy	2969	-----GGGTCAAATA	2969
Db	2969	-----GGGTCAAATA	2969

RESULT 10
US-10-982-512-20
Sequence 20, Application US/10982512
Publication No. US20050059081A1
GENERAL INFORMATION:
APPLICANT: Qi, Steve
APPLICANT: Akinsanya, Karen
APPLICANT: Riviere, Pierre
APPLICANT: Junion, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/10/982, 512
PRIOR FILING DATE: 2004-11-05
PRIOR APPLICATION NUMBER: US/09/976, 674
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
NUMBER OF SEQ ID NOS: 61
SOFTWARE: Patent in version 3.1

! SEQ ID NO 20
! LENGTH: 4676
! TYPE: DNA
! ORGANISM: Homo sapiens
US-10-982-512-20

Query Match 89.9%; Score 2806.4; DB 9; Length 4676;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;

Qy 1 AGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCTCTCTTTAGTGGCG 60
Db |||||
Qy 1 AGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCTCTCTTTAGTGGCG 60
Db |||||
Qy 61 CGTTCCGCCCTGGGTGTTCACCGCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCG 120
Db |||||
Qy 61 CGTTCCGCCCTGGGTGTTCACCGCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCG 120
Db |||||
Qy 121 GAGTGAGCGCGCGAGCATGAAAGCGCGCGAGCGCGCTCCATAGCGCACTGCGGACCG 180
Db |||||
Qy 121 GAGTGAGCGCGCGAGCATGAAAGCGCGCGAGCGCGCTCCATAGCGCACTGCGGACCG 180
Db |||||
Qy 181 TCCGCGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
Db |||||
Qy 181 TCCGCGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
Db |||||
Qy 241 CTGGGTGTTGAGATATTTCAAACTCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db |||||
Qy 241 CTGGGTGTTGAGATATTTGAAACTCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db |||||
Qy 301 CCTAAATTTGGAGCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAAAGCTGCTT 360
Db |||||
Qy 301 CCTAAATTTGGAGCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAAAGCTGCTT 360
Db |||||
Qy 361 GCCGATACAGAAATATCATGCTACATGATGGCTAGGACCAACACATGATTTTCATGTTT 420
Db |||||
Qy 361 GCCGATACAGAAATATCATGCTACATGATGGCTAGGACCAACACATGATTTTCATGTTT 420
Db |||||
Qy 421 GTCAAGAGGAATGATCCAGATGGAATCTATTGAGAGAAATCTATTACCTTGCATGCT 480
Db |||||
Qy 421 GTCAAGAGGAATGATCCAGATGGAATCTATTGAGAGAAATCTATTACCTTGCATGCT 480
Db |||||
Qy 481 GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTCCTCAAACTCAATAGAGCA 540
Db |||||
Qy 481 GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTCCTCAAACTCAATAGAGCA 540
Db |||||
Qy 541 GCAGTCTTAATGCTCTCTTGGAGAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db |||||
Qy 541 GCAGTCTTAATGCTCTCTTGGAGAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db |||||
Qy 601 GGAATGTATCTCGAAGAGAACTATTAAAGAGAAAGAAACCGATTGGAAACAGTCGGA 660
Db |||||
Qy 601 GGAATGTATCTCGAAGAGAACTATTAAAGAGAAAGAAACCGATTGGAAACAGTCGGA 660
Db |||||
Qy 661 ATTGCTCTTACGATTTATCCAAAGAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA 720
Db |||||
Qy 661 ATTGCTCTTACGATTTATCCAAAGAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA 720
Db |||||
Qy 721 ATTATACAGTAAAGATGGAGGCGCAAGGATTTTACGAACTCTTTAAAGGCCCAAT 780
Db |||||
Qy 721 ATTATACAGTAAAGATGGAGGCGCAAGGATTTTACGAACTCTTTAAAGGCCCAAT 780
Db |||||
Qy 781 CTAGTGGAACTAGTGTCTCCACATACGATGGATCCAAATTTATGCCCGCTGATCCA 840
Db |||||
Qy 781 CTAGTGGAACTAGTGTCTCCACATACGATGGATCCAAATTTATGCCCGCTGATCCA 840
Db |||||
Qy 841 GACTGGATTCCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAAACAGAGAA 900
Db |||||
Qy 841 GACTGGATTCCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAAACAGAGAA 900
Db |||||
Qy 901 GAAAGGAGACTCATTTATGTGCAATGAGCTAGCCAAATGGAGAGATGCCAGATCA 960
Db |||||
Qy 901 GAAAGGAGACTCATTTATGTGCAATGAGCTAGCCAAATGGAGAGATGCCAGATCA 960
Db |||||

Qy 961 GCTGGAGTGGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG 1020
Db |||||
Qy 961 GCTGGAGTGGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG 1020
Db |||||
Qy 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAATAATTTCTAGAAATTTCTATATGAAGAA 1080
Db |||||
Qy 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAATAATTTCTAGAAATTTCTATATGAAGAA 1080
Db |||||
Qy 1081 AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCCTATGTTGGAAACAGGAGG 1140
Db |||||
Qy 1081 AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCCTATGTTGGAAACAGGAGG 1140
Db |||||
Qy 1141 GCAGATTCATTCCGTTATCCTTAAACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAGATG 1200
Db |||||
Qy 1141 GCAGATTCATTCCGTTATCCTTAAACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAGATG 1200
Db |||||
Qy 1201 TCAGAAATATGATTCATGCTGAAGGAGGATCATAGATGTCATAGATAGAGAACTAAT 1260
Db |||||
Qy 1201 TCAGAAATATGATTCATGCTGAAGGAGGATCATAGATGTCATAGATAGAGAACTAAT 1260
Db |||||
Qy 1261 CAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAATATTTGCCAGAGCTGGATGACTCCT 1320
Db |||||
Qy 1261 CAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAATATTTGCCAGAGCTGGATGACTCCT 1320
Db |||||
Qy 1321 GAGGAAAAATATGCTTGGTCCATCTTACATAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Db |||||
Qy 1321 GAGGAAAAATATGCTTGGTCCATCTTACATAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Db |||||
Qy 1381 TTGATCTCACCTGAAATTTATTTCCAGTAGAAGATGATTTATGGAAGGAGGAGAGATC 1440
Db |||||
Qy 1381 TTGATCTCACCTGAAATTTATTTCCAGTAGAAGATGATTTATGGAAGGAGGAGAGATC 1440
Db |||||
Qy 1441 ATTGAGTCACTGCTGATCTGTGAGCGCACTAAATTTATCTATGAAGAAACAAACAGATC 1500
Db |||||
Qy 1441 ATTGAGTCACTGCTGATCTGTGAGCGCACTAAATTTATCTATGAAGAAACAAACAGATC 1500
Db |||||
Qy 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTCCCAAGTCCAGAAAGAGAAATTCAG 1560
Db |||||
Qy 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTCCCAAGTCCAGAAAGAGAAATTCAG 1560
Db |||||
Qy 1561 TTTATTTTTCCTCTGATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620
Db |||||
Qy 1561 TTTATTTTTCCTCTGATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620
Db |||||
Qy 1621 TTTAAAGAAACAGAAATATAAATCCAGTCCAGTGGTGGCTGCTGCTCCAAAGTGAATTCAG 1680
Db |||||
Qy 1621 TTTAAAGAAACAGAAATATAAATCCAGTCCAGTGGTGGCTGCTGCTCCAAAGTGAATTCAG 1680
Db |||||
Qy 1681 TGTCTTATCAAAGAGGAGATAGCAATTTACAGTGGTGAATGGAAATTTCTTGGCGCGCAT 1740
Db |||||
Qy 1681 TGTCTTATCAAAGAGGAGATAGCAATTTACAGTGGTGAATGGAAATTTCTTGGCGCGCAT 1740
Db |||||
Qy 1741 GGATCTAATATCCAAAGTGGATGAATGAGTCAAGAGCTGGTATATTTTGAAGGACCCAAAGAC 1800
Db |||||
Qy 1741 GGATCTAATATCCAAAGTGGATGAATGAGTCAAGAGCTGGTATATTTTGAAGGACCCAAAGAC 1800
Db |||||
Qy 1801 TCCCTTTTAGAGCATCACCTGTAGTACGTTACGTTAAATCCCTGGAGAGGTGACAAGG 1860
Db |||||
Qy 1801 TCCCTTTTAGAGCATCACCTGTAGTACGTTAAATCCCTGGAGAGGTGACAAGG 1860
Db |||||
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCACTGTGACTTCTTTTATA 1920
Db |||||
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCACTGTGACTTCTTTTATA 1920
Db |||||
Qy 1921 AGTAAGTATATGATACAGAGAAATCCACTGTGTGCTCCCTTTACAGCTATCAAGTCTCT 1980
Db |||||
Qy 1921 AGTAAGTATATGATACAGAGAAATCCACTGTGTGCTCCCTTTACAGCTATCAAGTCTCT 1980
Db |||||
Qy 1981 GAAAGTACCCCACTTCGAAACAAAGGAATTTTGGGCCCACTTTTGGATTCACAGCT 2040
Db |||||
Qy 1981 GAAAGTACCCCACTTCGAAACAAAGGAATTTTGGGCCCACTTTTGGATTCACAGCT 2040
Db |||||

Qy 2041 CCTCTCTCGACTATACCTCTCGAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 2100
Db 2041 CCTCTCTCGACTATACCTCTCGAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 2100
Qy 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
Qy 2161 CTGTTTCATATATGGTGGTCTCTCAGGTGCAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTTCATATATGGTGGTCTCTCAGGTGCAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
Qy 2221 TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTGTTGTAGTGATAGACACAGG 2280
Db 2180 ----- 2179
Qy 2281 GGATCCTGTCCAGGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAATA 2340
Db 2180 -----GGGTCAATA 2189
Qy 2341 GAAATTGAGATCAGGTGGAGGACTCCATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2190 GAAATTGAGATCAGGTGGAGGACTCCATATCTAGCTTCTCGATATGATTTCAATTGAC 2249
Qy 2401 TTAGATCGTGTGGGATCCAGCGTGTCTTATGGAGGATACCTCTCCCTGATGGCATTA 2460
Db 2250 TTAGATCGTGTGGGATCCAGCGTGTCTTATGGAGGATACCTCTCCCTGATGGCATTA 2309
Qy 2461 ATGCAAGGTCAAGATATCTTCAAGGGTTGCTTATGCTGGGGCCCACTCATCTCTGTGGATC 2520
Db 2310 ATGCAAGGTCAAGATATCTTCAAGGGTTGCTTATGCTGGGGCCCACTCATCTCTGTGGATC 2369
Qy 2521 TTCTATGATACAGGATACACGGAACTGTTATATGGGTCACTCCCTCAGCAGATGAACAGGGC 2580
Db 2370 TTCTATGATACAGGATACACGGAACTGTTATATGGGTCACTCCCTCAGCAGATGAACAGGGC 2429
Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Db 2430 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2489
Qy 2641 CTGCTCTTACATGGTTTCTTGATGAGAACTGCAATTTTGCAATACCACTGATTTACTG 2700
Db 2490 CTGCTCTTACATGGTTTCTTGATGAGAACTGCAATTTTGCAATACCACTGATTTACTG 2549
Qy 2701 AGTATTAAGTCTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACAC 2760
Db 2550 AGTATTAAGTCTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACAC 2609
Qy 2761 AGCATTAAGTCTGATTCGGAAGCAATTAATCAATGATCTTTTGCACTACCTTCAA 2820
Db 2610 AGCATTAAGTCTGATTCGGAAGCAATTAATGATCTTTTGCACTACCTTCAA 2669
Qy 2821 GAAAACTTGGATCAGTATTTGCTCTCTAAAGTGATTAATTTGACCTGTGAGAAC 2880
Db 2670 GAAAACTTGGATCAGTATTTGCTCTCTAAAGTGATTAATTTGACCTGTGAGAAC 2729
Qy 2881 TCTCTGGTATACATGGCTATTATAACCAATGAGAGGTTTAATCAACAGAAAACACAGA 2940
Db 2730 TCTCTGGTATACATGGCTATTATAACCAATGAGAGGTTTAATCAACAGAAAACACAGA 2789
Qy 2941 ATTGATCATCATATTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAATGGTGG 3000
Db 2790 ATTGATCATCATATTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAATGGTGG 2849
Qy 3001 CCATGAGGGGTCTACGGTTTGTGGTAGTAACTTAATACCTTAAACCCCATGCTCAAAA 3060
Db 2850 CCATGAGGGGTCTACGGTTTGTGGTAGTAACTTAATACCTTAAACCCCATGCTCAAAA 2909
Qy 3061 TCAAATGATACATATTTCTGAGAGACCCGAGCAATACCATAGAAATTAATAAAAAAAA 3120
Db 2910 TCAAATGATACATATTTCTGAGAGACCCGAGCAATACCATAGAAATTAATAAAAAAAA 2969

RESULT 11
US-10-170-789-39
; Sequence 39, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 2649
; TYPE: DNA

i ORGANISM: Homo sapiens
us-10-170-789-39

Query Match 84.8%; Score 2645.8; DB 6; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	214	ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTGAGATATTTGAAACCTGCGACTGT	273
Db	1	ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTGAGATATTTGAAACCTGCGACTGT	60
QY	274	GAGGAGAATATTGGAATCACAGGATCGGCCCTAAATTTGGAGCCTTTTATTTGTTGAGCGGTAT	333
Db	61	GAGGAGAATATTGGAATCACAGGATCGGCCCTAAATTTGGAGCCTTTTATTTGTTGAGCGGTAT	120
QY	334	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGGATACAGGAATATCATGGCTACATGATG	393
Db	121	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGGATACAGGAATATCATGGCTACATGATG	180
QY	394	GCTAAGCACCACATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTTCA	453
Db	181	GCTAAGGACCACATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTTCA	240
QY	454	GACAGAATCTATTACCTTGGCATGCTCTGGTGAACACAGAGAAATACACTGTTTATTTCT	513
Db	241	GACAGAATCTATTACCTTGGCATGCTCTGGTGAACACAGAGAAATACACTGTTTATTTCT	300
QY	514	GAATTTCCAAAACATATCAATAGAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTTG	573
Db	301	GAATTTCCAAAACATATCAATAGAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTTG	360
QY	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTTCGAGAAGAACTATTAAAGA	633
Db	361	GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTTCGAGAAGAACTATTAAAGA	420
QY	634	GAAGAAGAACCATTTGGAACAGCTCGGAATTTGCTTTTACGATTTATCACCAGAGAAAGTGA	693
Db	421	GAAGAAGAACCATTTGGAACAGCTCGGAATTTGCTTTTACGATTTATCACCAGAGAAAGTGA	480
QY	694	ACATTTCTGTTTCAAGCCGGTGTGGAATTTATCAAGTAAAGATGGAGGGCCACAAGGA	753
Db	481	ACATTTCTGTTTCAAGCCGGTGTGGAATTTATCAAGTAAAGATGGAGGGCCACAAGGA	540
QY	754	TTTACGGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCAACTACGGATG	813
Db	541	TTTACGGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCAACTACGGATG	600
QY	814	GATCCAAATTTATGCCCTGCTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATAT	873
Db	601	GATCCAAATTTATGCCCTGCTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATAT	660
QY	874	TGGATATCTAAACATCGTAACACAGAGAGAAAGAGAGACTCATCTTATGTGCAATGAGCTA	933
Db	661	TGGATATCTAAACATCGTAACACAGAGAGAAAGAGAGACTCATCTTATGTGCAATGAGCTA	720
QY	934	GCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTGCCTTACCTTTGTTCTCCAGAGAA	993
Db	721	GCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTGCCTTACCTTTGTTCTCCAGAGAA	780
QY	994	TTTGTAGATATTCTGGCTATTGTTGGTGTCCAAAAGCTGAACAACTCCAGTGTGTGT	1053
Db	781	TTTGTAGATATTCTGGCTATTGTTGGTGTCCAAAAGCTGAACAACTCCAGTGTGTGT	840
QY	1054	AAAAATTTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113
Db	841	AAAAATTTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	900
QY	1114	ACATCCCTCTATGTTGGAAACAAAGGAGGCGAGATTCATTCGGTTATCTCTAAACAGGTACA	1173
Db	901	ACATCCCTCTATGTTGGAAACAAAGGAGGCGAGATTCATTCGGTTATCTCTAAACAGGTACA	960
QY	1174	GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGAGGAAGGATC	1233

Db	961	GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGCTCGAAGGAAGATC	1020
QY	1234	ATAGATGTCTATAGATAAGGAACTAAATTCACCTTTTGGATGATCTTATTTGAAGAGTTGAA	1293
Db	1021	ATAGATGTCTATAGATAAGGAACTAAATTCACCTTTTGGATGATCTTATTTGAAGAGTTGAA	1080
QY	1294	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATATGCTTGGTCCATCTCTACTAGAT	1353
Db	1081	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATATGCTTGGTCCATCTCTACTAGAT	1140
QY	1354	CGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA	1413
Db	1141	CGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA	1200
QY	1414	GATGATGTTTATGGAAGGCGAGAGACTCAITGAGTCAAGTGTGCTGATCTGTGAGCGCACCTA	1473
Db	1201	GATGATGTTTATGGAAGGCGAGAGACTCAITGAGTCAAGTGTGCTGATCTGTGAGCGCACCTA	1260
QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATGAATATCCATCAGATCTTTTCAATGTTTTT	1533
Db	1261	ATTATCTATGAAGAAACACAGACATCTGGATGAATATCCATCAGATCTTTTCAATGTTTTT	1320
QY	1534	CCCCAAGTCAAGAGGGAATTTGAGTTTATTTTTCCTCTGATGTCGCAAAACAGGTTTC	1593
Db	1321	CCCCAAGTCAAGAGGGAATTTGAGTTTATTTTTCCTCTGATGTCGCAAAACAGGTTTC	1380
QY	1594	CGTCAATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAAAGATCAAGTGGT	1653
Db	1381	CGTCAATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAAAGATCAAGTGGT	1440
QY	1654	GGGCTGCCTGCTCCAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGCAATTTACCGT	1713
Db	1441	GGGCTGCCTGCTCCAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGCAATTTACCGT	1500
QY	1714	GGTGAATGGGAAGTTCTTTGGCGGCATGATCTAATATCCAAGTTGATGAAGTCAGAAG	1773
Db	1501	GGTGAATGGGAAGTTCTTTGGCGGCATGATCTAATATCCAAGTTGATGAAGTCAGAAG	1560
QY	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833
Db	1561	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1620
QY	1834	TACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGC	1893
Db	1621	TACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGC	1680
QY	1894	ATCAGTCAAGCTGTGACTTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGT	1953
Db	1681	ATCAGTCAAGCTGTGACTTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGT	1740
QY	1954	GTGTCCTTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAACAAAGGAATTT	2013
Db	1741	GTGTCCTTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAACAAAGGAATTT	1800
QY	2014	TGGGCCACCAATTTTGGATTTGAGCAGGTCTCTCTCTGATATATCTCTCCAGAAAATTTTC	2073
Db	1801	TGGGCCACCAATTTTGGATTTGAGCAGGTCTCTCTCTGATATATCTCTCCAGAAAATTTTC	1860
QY	2074	TCCTTTGAAAGTACTACTGGAATTTACATTTGATGGAGTCTCTACAGCCCTCATGATCTA	2133
Db	1861	TCCTTTGAAAGTACTACTGGAATTTACATTTGATGGAGTCTCTACAGCCCTCATGATCTA	1920
QY	2134	CAGCTTGGAAAGAAATATCTCTGCTGTTTCAATATGTTGTTGCTCTCAGGTGAGTTG	2193
Db	1921	CAGCTTGGAAAGAAATATCTCTGCTGTTTCAATATGTTGTTGCTCTCAGGTGAGTTG	1980
QY	2194	GTGAATAATCGGTTTTAAAGAGTCAAGTATTTTCCCTTTGAATACCTTAGCCTCTCTAGGT	2253
Db	1981	GTGAATAATCGGTTTTAAAGAGTCAAGTATTTTCCGCTTGAATACCTTAGCCTCTCTAGGT	2040
QY	2254	TATGTGGTTGTAGTATAGACAAACAGGGGATCTCTGTCCAGGGGCTTTAAATTTGAAGGC	2313
Db	2041	TATGTGGTTGTAGTATAGACAAACAGGGGATCTCTGTCCAGGGGCTTTAAATTTGAAGGC	2100

Qy	2314	GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAAATAT	2373
Db	2101	GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAAATAT	2160
Qy	2374	CTAGCTTCTCGATATGATTTTCATTTGATCTAGATCGTGGGATCCACGGCTGGTCCCTAT	2433
Db	2161	CTAGCTTCTCGATATGATTTTCATTTGATCTAGATCGTGGGATCCACGGCTGGTCCCTAT	2220
Qy	2434	GGAGGATACCTTCCCTGATGGCATTAAATCCAGAGTCAAGATATCTTCAGGGTTGCTATT	2493
Db	2221	GGAGGATACCTTCCCTGATGGCATTAAATCCAGAGTCAAGATATCTTCAGGGTTGCTATT	2280
Qy	2494	GCTGGGGCCCCAGTCACTCTGTGTGATCTTCTATGATACAGGATACACGGAACTTTATATG	2553
Db	2281	GCTGGGGCCCCAGTCACTCTGTGTGATCTTCTATGATACAGGATACACGGAACTTTATATG	2340
Qy	2554	GGTCAACCTCGACCAAGATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGAA	2613
Db	2341	GGTCAACCTCGACCAAGATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGAA	2400
Qy	2614	AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAAATGTC	2673
Db	2401	AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAAATGTC	2460
Qy	2674	CATTTTGCACATACACGATATATTACTGAGTTTTTTTAACTGAGGCTGGAAAGCCATATGAT	2733
Db	2461	CATTTTGCACATACACGATATATTACTGAGTTTTTTTAACTGAGGCTGGAAAGCCATATGAT	2520
Qy	2734	TTACAGATCTATCTCTCAGGAGACACAGCATAAAGATTTCCTGAAATCGGAGAACATTAT	2793
Db	2521	TTACAGATCTATCTCTCAGGAGACACAGCATAAAGATTTCCTGAAATCGGAGAACATTAT	2580
Qy	2794	GAATGATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTTAAAA	2853
Db	2581	GAATGATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTTAAAA	2640
Qy	2854	GTGATATNA 2862	
Db	2641	GTGATATNA 2649	

RESULT 12
US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-2

Query Match 84.8%; Score 2644.4; DB 3; Length 2671;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	209	GCAATGCGCAGCAGCAATGGAACAGACAGCTGGGTGGATATTTGAACTGCGG	268
Db	3	GTACATGCGCAGCAGCAATGGAACAGACAGCTGGGTGGATATTTGAACTGCGG	62

Qy	269	ACTGTGAGGAGAAATTTGAATACAGGATCGCCTAAATTTGAGACCTTTTATGTTGAGC	328
Db	63	ACTGTGAGGAGAAATTTGAATACAGGATCGCCTAAATTTGAGACCTTTTATGTTGAGC	122
Qy	329	GGTATTTCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACAGAAATATCATGGCTACA	388
Db	123	GGTATTTCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACAGAAATATCATGGCTACA	182
Qy	389	TGATGGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTC	448
Db	183	TGATGGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTC	242
Qy	449	ATTTCAGACAGAAATCTATTACCTTGCCTATGCTGTGTGAGAAACAGAGAAATATACATGTTTT	508
Db	243	ATTTCAGACAGAAATCTATTACCTTGCCTATGCTGTGTGAGAAACAGAGAAATATACATGTTTT	302
Qy	509	ATTTCTGAAATTTCCCAAACTATCAATAGACAGCAGCTTTAAATGCTCTCTTGGAAAGCCTC	568
Db	303	ATTTCTGAAATTTCCCAAACTATCAATAGACAGCAGCTTTAAATGCTCTCTTGGAAAGCCTC	362
Qy	569	TTTTTGGATCTTTTTTTCAGGCAACACTGGAATGGAATGATTTCTTCGAGAGAAAGAACTAT	628
Db	363	TTTTTGGATCTTTTTTTCAGGCAACACTGGAATGGAATGATTTCTTCGAGAGAAAGAACTAT	422
Qy	629	TAAAGAGAAAGAAACCGCATTCGGAATTCGGAATTCGCTTCTTACGATTTATCACAAGGAA	688
Db	423	TAAAGAGAAAGAAACCGCATTCGGAATTCGGAATTCGCTTCTTACGATTTATCACAAGGAA	482
Qy	689	GTGGAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAGATGGAGGCCAC	748
Db	483	GTGGAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAGATGGAGGCCAC	542
Qy	749	AAGGATTTTACCCCAACACCTTTTAAAGCCCAATCTAGTGGAAACCTAGTTGTCCTCAACATAC	808
Db	543	AAGGATTTTACCCCAACACCTTTTAAAGCCCAATCTAGTGGAAACCTAGTTGTCCTCAACATAC	602
Qy	809	GGATGGATCCAAATTTATGCTGCTGATCCAGCTGGATTCGTTTATACATAGCAAG	868
Db	603	GGATGGATCCAAATTTATGCTGCTGATCCAGCTGGATTCGTTTATACATAGCAAG	662
Qy	869	ATATTTGGATCTTAAACATCGTAAACAGAGAAAGGAGACTCCTATTTGTCACAAATG	928
Db	663	ATATTTGGATCTTAAACATCGTAAACAGAGAAAGGAGACTCCTATTTGTCACAAATG	722
Qy	929	AGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTTCTTGTCTTCTCAAG	988
Db	723	AGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTTCTTGTCTTCTCAAG	782
Qy	989	AAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACAACTCCAGTG	1048
Db	783	AAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACAACTCCAGTG	842
Qy	1049	GTGGTAAAAATCTTAGAATTTCTATGAAGAAATATGATGATCTGAGGTGGAATTTATTC	1108
Db	843	GTGGTAAAAATCTTAGAATTTCTATGAAGAAATATGATGATCTGAGGTGGAATTTATTC	902
Qy	1109	ATGTTTACATCCCTATGTTGGAAACAAAGGAGGAGAGATTCATTCCTGTTATCTTAAACAG	1168
Db	903	ATGTTTACATCCCTATGTTGGAAACAAAGGAGGAGAGATTCATTCCTGTTATCTTAAACAG	962
Qy	1169	GTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGTCGTAAGGAA	1228
Db	963	GTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGTCGTAAGGAA	1022
Qy	1229	GGATCATAGATGTCATAGATGAAGAACTTAATTCACCTTTTGGATTTCTTATGAGGAG	1288
Db	1023	GGATCATAGATGTCATAGATGAAGAACTTAATTCACCTTTTGGATTTCTTATGAGGAG	1082
Qy	1289	TTGAATATATTTCCGAGAGCTGGATGGAATCTGAGGGGAAATATGCTGGTCCATCTAC	1348
Db	1083	TTGAATATATTTCCGAGAGCTGGATGGAATCTGAGGGGAAATATGCTGGTCCATCTAC	1142
Qy	1349	TAGATCGCTCCGAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAG	1408

1143 TAGATCGCTCCAGAGCTCGCTGAGATAGTGTGATCTCACTGATTTATTTATCCAG 1202
1409 TAGAAGATGATGTTATGAAAGGAGAGAGCTCAATGAGTCAGTGCCTGATTTCTGAGCGC 1468
1203 TAGAAGATGATGTTATGAAAGGAGAGAGCTCAATGAGTCAGTGCCTGATTTCTGAGCGC 1262
1469 CACTAATTTATCTATGAAAGAACACAGACATCTGGATAAATATCCATGATCTTTTCATG 1528
1263 CACTAATTTATCTATGAAAGAACACAGACATCTGGATAAATATCCATGATCTTTTCATG 1322
1529 TTTTTCCTCCAAAGTCAGAGAGAAATTCAGTTTATTTTGGCTCTGAAATCAAAACAG 1588
1323 TTTTTCCTCCAAAGTCAGAGAGAAATTCAGTTTATTTTGGCTCTGAAATCAAAACAG 1382
1589 GTTTCCTGTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAAAATATAACGATCCA 1648
1383 GTTTCCTGTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAAAATATAACGATCCA 1442
1649 GTGTTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTATCAAAGAGGAGATAGCAATTA 1708
1443 GTGTTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTATCAAAGAGGAGATAGCAATTA 1502
1709 CCAGTGGTGAATGGAGTCTTTCGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCA 1768
1503 CCAGTGGTGAATGGAGTCTTTCGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCA 1562
1769 GAAGGCTGGTATTTTGAAGGCACCAAGATCTCCCTTTAGAGCATCACTGTAGTAG 1828
1563 GAAGGCTGGTATTTTGAAGGCACCAAGATCTCCCTTTAGAGCATCACTGTAGTAG 1622
1829 TCAGTTAGCTAAATCTGGAGAGGTGAACAAGGCTGACTGACCTGGCTACTCACATCTTT 1888
1623 TCAGTTAGCTAAATCTGGAGAGGTGAACAAGGCTGACTGACCTGGCTACTCACATCTTT 1682
1889 GCTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAAGTATAGTAAACCAAGAAATCCAC 1948
1683 GCTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAAGTATAGTAAACCAAGAAATCCAC 1742
1949 ACTGTGTGCTCCTTTACAAAGTATCAAGTCTGAAAGTACCAAGTCTGAAACCAAGG 2008
1743 ACTGTGTGCTCCTTTACAAAGTATCAAGTCTGAAAGTACCAAGTCTGAAACCAAGG 1802
2009 AATTTTGGGCCACCAATTTGGATTCAGCAGGTCTCTTCTGACTACTACTCTCCAGAAA 2068
1803 AATTTTGGGCCACCAATTTGGATTCAGCAGGTCTCTTCTGACTACTACTCTCCAGAAA 1862
2069 TTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAGCCTCATG 2128
1863 TTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAGCCTCATG 1922
2129 ATCTACAGCTGGAAGAAATATCTTACTGTGCTGTTTATATATGTTGGTCTTCAGGTGC 2188
1923 ATCTACAGCTGGAAGAAATATCTTACTGTGCTGTTTATATATGTTGGTCTTCAGGTGC 1982
2189 AGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTC 2248
1983 AGTGGTGAATTAATCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTC 2042
2249 TAGGTTATGTTGTTGATGATAGACAAAGGGGATCTGTCTACCGAGGGCTTAAATTTG 2308
2043 TAGGTTATGTTGTTGATGATAGACAAAGGGGATCTGTCTACCGAGGGCTTAAATTTG 2102
2309 AAGGGCTCTTTAAATATATAAATGGTCAAAATAGAGATGAGTCAAGTGGAGGACTCC 2368
2103 AAGGGCTCTTTAAATATATAAATGGTCAAAATAGAGATGAGTCAAGTGGAGGACTCC 2162
2369 AATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCAGGCTGTG 2428
2163 AATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCAGGCTGTG 2222
2429 CCTATGGAGATACCTCTCCCTGATGGCATTTAATGACAGAGTCAGATATCTTCAGGGTTG 2488

2223 CCTATGGAGGATACCTCTCCCTGATGGCATTTAATGACAGAGTCAGATATCTTCAGGGTTG 2282
2489 CTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTT 2548
2283 CTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTT 2342
2549 ATATGGGTCACTCCCTGACAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAG 2608
2343 ATATGGGTCACTCCCTGACAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAG 2402
2609 CAGAAAGTTCCCTCTCTGAAACCAATCGTTTACTGTCTTTACATGGTTTCTTGGATGAGA 2668
2403 CAGAAAGTTCCCTCTCTGAAACCAATCGTTTACTGTCTTTACATGGTTTCTTGGATGAGA 2462
2669 ATGTCCATTTTGCATACATACAGTATATTACTAGTTTATTTAGTACAGGCTGGAAGCCAT 2728
2463 ATGTCCATTTTGCATACATACAGTATATTACTAGTTTATTTAGTACAGGCTGGAAGCCAT 2522
2729 ATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGTTCCTGAAATCGGGAGAAC 2788
2523 ATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGTTCCTGAAATCGGGAGAAC 2582
2789 ATTATGAATCGATCTTTTGCATCTTCTGCACTACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTC 2848
2583 ATTATGAATCGATCTTTTGCATCTTCTGCACTACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTC 2642
2849 TAAAGTGATATAA 2862
2643 TAAAGTGATATAA 2656

RESULT 13
US-10-982-512-2
; Sequence 2, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-2

Query Match 84.8%; Score 2644.4; DB 9; Length 2671;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 209 GCACATGGCAGCAGCAATGGAAAACAGAACAGCTGGTGTTCAGATATTTGAAATCGCGG 268
DB 3 GTACCATGGCAGCAGCAATGGAAAACAGAACAGCTGGTGTTCAGATATTTGAAATCGCGG 62
QY 269 ACTGTGAGGAGATATTGAAATCAGAGATCGGCCTTAAATTTGGAGCTTTTATGTTGAGC 328
DB 63 ACTGTGAGGAGATATTGAAATCAGAGATCGGCCTTAAATTTGGAGCTTTTATGTTGAGC 122
QY 329 GGTATTCTCGGAGTCAGCTTAAAGAGCTGTTCCGATACCCAGAAAATATCATGGCTACA 388
DB 123 GGTATTCTCGGAGTCAGCTTAAAGAGCTGTTCCGATACCCAGAAAATATCATGGCTACA 182
QY 389 TGATGGCTAAGGCCACCAATGATTTTCAATGTTGTGAAGGAGATGATCCAGATGGACCTC 448

[illegible]

Db	1263	CAC	TAA	TAT	CTA	TGA	AGAA	CAAC	AAGA	CAT	CTGG	ATAA	TAT	TCA	TGA	CAT	CTT	CTC	ATG	1322	
Qy	1529	TTTT	CCCC	AAAG	TCA	GGA	AGAA	TTAG	TTAT	TTTT	TG	CC	TCT	GAA	TG	CAAAA	CAG		1588		
Db	1323	TTTT	TT	CCCC	AAAG	TCA	GGA	AGAA	TTAG	TTAT	TTTT	TG	CC	TCT	GAA	TG	CAAAA	CAG	1382		
Qy	1589	GT	TT	CCG	TCA	TTT	A	CAAAA	TAC	AT	CT	TTTT	AA	AGAA	GCAAA	TAT	AAAC	GCAT	CCA	1648	
Db	1383	GT	TT	CCG	TCA	TTT	A	CAAAA	TAC	AT	CT	TTTT	AA	AGAA	GCAAA	TAT	AAAC	GCAT	CCA	1442	
Qy	1649	GT	GG	GG	GC	TCC	TCT	CAAG	TGA	TTT	CAA	GT	GT	CCT	AT	CAA	AG	GAG	GAT	AG	1708
Db	1443	GT	GG	GG	GC	TCC	TCT	CAAG	TGA	TTT	CAA	GT	GT	CCT	AT	CAA	AG	GAG	GAT	AG	1502
Qy	1709	CCA	G	TG	TG	GA	TT	CGG	CGG	CAT	GGAT	CT	TA	AT	AT	CCA	AG	TTC	GAT	CA	1768
Db	1503	CCA	G	TG	TG	GA	TT	CGG	CGG	CAT	GGAT	CT	TA	AT	AT	CCA	AG	TTC	GAT	CA	1562
Qy	1769	GAA	G	CT	GG	TAT	TTTT	GAA	GG	CA	CAAA	GAC	T	CCCC	TTT	AG	A	G	CAT	C	1828
Db	1563	GAA	G	CT	GG	TAT	TTTT	GAA	GG	CA	CAAA	GAC	T	CCCC	TTT	AG	A	G	CAT	C	1622
Qy	1829	TC	AG	T	AC	GT	AA	T	CT	CG	AG	AG	T	CA	AG	CT	CA	CT	CA	CT	1888
Db	1623	TC	AG	T	AC	GT	AA	T	CT	CG	AG	AG	T	CA	AG	CT	CA	CT	CA	CT	1682
Qy	1889	G	CT	CA	CT	CAG	C	AC	T	GT	CA	CT	T	TT	T	AA	G	T	AG	T	1948
Db	1683	G	CT	CA	CT	CAG	C	AC	T	GT	CA	CT	T	TT	T	AA	G	T	AG	T	1742
Qy	1949	A	CT	GT	GT	CC	TTT	A	CA	AG	CT	CT	G	A	G	A	T	G	A	G	2008
Db	1743	A	CT	GT	GT	CC	TTT	A	CA	AG	CT	CT	G	A	G	A	T	G	A	G	1802
Qy	2009	AA	TTT	T	G	GG	CA	CA	AT	TTT	T	G	GA	T	CA	G	AG	GT	C	CT	2068
Db	1803	AA	TTT	T	G	GG	CA	CA	AT	TTT	T	G	GA	T	CA	G	AG	GT	C	CT	1862
Qy	2069	TTTT	CT	CT	TTTT	G	AA	AG	T	CA	CT	GG	AT	TT	CA	AT	GT	T	G	GA	2128
Db	1863	TTTT	CT	CT	TTTT	G	AA	AG	T	CA	CT	GG	AT	TT	CA	AT	GT	T	G	GA	1922
Qy	2129	A	T	C	A	G	C	T	G	G	A	A	A	A	T	C	C	T	A	T	2188
Db	1923	A	T	C	A	G	C	T	G	G	A	A	A	A	T	C	C	T	A	T	1982
Qy	2189	A	G	T	CG	T	CA	A	A	T	CG	TT	AA	A	GG	AG	T	CA	AG	T	2248
Db	1983	A	G	T	CG	T	CA	A	A	T	CG	TT	AA	A	GG	AG	T	CA	AG	T	2042
Qy	2249	TAG	TT	A	T	G	CG	TT	A	G	T	AG	T	CA	A	CA	A	G	GG	AT	2308
Db	2043	TAG	TT	A	T	G	CG	TT	A	G	T	AG	T	CA	A	CA	A	G	GG	AT	2102
Qy	2309	A	AG	CG	CC	TT	AA	A	T	AA	A	GG	T	CA	A	T	GA	AA	T	CA	2368
Db	2103	A	AG	CG	CC	TT	AA	A	T	AA	A	GG	T	CA	A	T	GA	AA	T	CA	2162
Qy	2369	A	A	T	A	T	C	A	G	T	T	C	A	T	G	A	T	T	C	A	2428
Db	2163	A	A	T	A	T	C	A	G	T	T	C	A	T	G	A	T	T	C	A	2222
Qy	2429	C	C	T	A	G	G	A	T	A	A	T	G	C	A	T	A	T	A	T	2488
Db	2223	C	C	T	A	G	G	A	T	A	A	T	G	C	A	T	A	T	A	T	2282
Qy	2489	C	T	A	T	G	C	T	G	G	G	C	C	C	A	G	T	C	A	T	2548
Db	2283	C	T	A	T	G	C	T	G	G	G	C	C	C	A	G	T	C	A	T	2342
Qy	2549	A	T	A	G	G	G	T	C	A	C	C	A	A	T	G	A	A	T	A	2608
Db	2343	A	T	A	G	G	G	T	C	A	C	C	A	A	T	G	A	A	T	A	2402

QY 2609 CAGAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGA 2668
DB 2403 CAGAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGA 2462
QY 2669 ATGTCCATTTTGCACATACAGTATATTTACTGAGTTTATTTAGTGGGCTGGAAAGCCAT 2728
DB 2463 ATGTCCATTTTGCACATACAGTATATTTACTGAGTTTATTTAGTGGGCTGGAAAGCCAT 2522
QY 2729 ATGATTTTACAGATCTATCTCTCAGGAGAGACACAGATTAAGAGTTTCTGAAATCGGAGAAC 2788
DB 2523 ATGATTTTACAGATCTATCTCTCAGGAGAGACACAGATTAAGAGTTTCTGAAATCGGAGAAC 2582
QY 2789 ATTATGAATCGCATCTTTTGGCACTACCTTCAAGAAAACCTTTGGATACGATTTGCTGCTC 2848
DB 2583 ATTATGAATCGCATCTTTTGGCACTACCTTCAAGAAAACCTTTGGATACGATTTGCTGCTC 2642
QY 2849 TAAAAGTGATATAA 2862
DB 2643 TAAAAGTGATATGA 2656

RESULT 14

US-10-054-776-1
; Sequence 1, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2649)
US-10-054-776-1

Query Match 84.7%; Score 2644.2; DB 6; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2646; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 273
DB 1 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 60
QY 274 GAGAGAAATTTGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATGTTGACGGTAT 333
DB 61 GAGGAGAATATTGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATGTTGACGGTAT 120
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACCGAATAATATCATGGCTACATCATG 393
DB 121 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACCGAATAATATCATGGCTACATCATG 180
QY 394 GCTAAGCACCACATGATTTTATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
DB 181 GCTAAGGACCAACATGATTTTATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 240
QY 454 GACAGAAATCTATTACCTTGGCCATGCTGTTGGTGAACACAGAAAATACACTGTTTTATCT 513
DB 241 GACAGAAATCTATTACCTTGGCCATGCTGTTGGTGAACACAGAAAATACACTGTTTTATCT 300
QY 514 GAAATTCGAAAACATATCAATAGACGAGCTGTTAATGCTCTTGGAGCCCTTTTGG 573
DB 301 GAAATTCGAAAACATATCAATAGACGAGCTTAAATGCTCTTGGAGCCCTTTTGG 360
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTTCGAGAGAGAACTATTAAGA 633

DB 361 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAACTATTAAGA 420
QY 634 GAAAGAAAACGCAATTTGGAAACAGTTCGGAATTTGCTTTACGATTTATCAACCAAGGAAGTGA 693
DB 421 GAAAGAAAACGCAATTTGGAAACAGTTCGGAATTTGCTTTACGATTTATCAACCAAGGAAGTGA 480
QY 694 ACATTTCTGTTTCAAGCCGTTAGTGAATTTATACGTTAAAGATGGAAGGCGCCACAGGA 753
DB 481 ACATTTCTGTTTCAAGCCGTTAGTGAATTTATACGTTAAAGATGGAAGGCGCCACAGGA 540
QY 754 TTTACGCAACCAACCTTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCCAACATACGATG 813
DB 541 TTTACGCAACCAACCTTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCCAACATACGATG 600
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATATACATAGCAACGATTT 873
DB 601 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATATACATAGCAACGATTT 660
QY 874 TGGATATCTAAACATCGTAAACAGAGAGAAAGGAGACTCCTTTATGTGCACAAATGAGCTA 933
DB 661 TGGATATCTAAACATCGTAAACAGAGAGAAAGGAGACTCCTTTATGTGCACAAATGAGCTA 720
QY 934 GCCAACAATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA 993
DB 721 GCCAACAATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA 780
QY 994 TTTGATAGATATTTCTGGCTATTTGGTGTCTCANAAGCTGAAACAACCTCCAGTGGTGT 1053
DB 781 TTTGATAGATATTTCTGGCTATTTGGTGTCTCANAAGCTGAAACAACCTCCAGTGGTGT 840
QY 1054 AAAATCTTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGGAAAATTTATCATGTT 1113
DB 841 AAAATCTTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGGAAAATTTATCATGTT 900
QY 1114 ACATCCCTATGTTGGAAACAGAGGCGAGATTTCAATCCGTTATCTTAAACAGAGTACA 1173
DB 901 ACATCCCTATGTTGGAAACAGAGGCGAGATTTCAATCCGTTATCTTAAACAGAGTACA 960
QY 1174 GCAATCTTAAAGTCACTTTTAAAGATGTCAGAAAATAATGATGATGCTGAAGAGGATC 1233
DB 961 GCAATCTTAAAGTCACTTTTAAAGATGTCAGAAAATAATGATGATGCTGAAGAGGATC 1020
QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAATTTCTATTTGAAGAGGATGAA 1293
DB 1021 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAATTTCTATTTGAAGAGGATGAA 1080
QY 1294 TATATGCCAGAGCTGGATGGACTCTGAGGAAAATAATGCTTGGTCCATCTACTAGAT 1353
DB 1081 TATATGCCAGAGCTGGATGGACTCTGAGGAAAATAATGCTTGGTCCATCTACTAGAT 1140
QY 1354 CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAGTAGAA 1413
DB 1141 CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAGTAGAA 1200
QY 1414 GATGATGTTATGGAAGGCGAGACTCAATTTGAGTCACTGCTGATTTGTGACGCCACTA 1473
DB 1201 GATGATGTTATGGAAGGCGAGACTCAATTTGAGTCACTGCTGATTTGTGACGCCACTA 1260
QY 1474 ATATATCTATGAAGAAACACAGACATCTCGATTAATATTCATGACATCTTTTCATGTTTTT 1533
DB 1261 ATATATCTATGAAGAAACACAGACATCTCGATTAATATTCATGACATCTTTTCATGTTTTT 1320
QY 1534 CCCCAAGTCCAGAGGAAAATTTGATTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
DB 1321 CCCCAAGTCCAGAGGAAAATTTGATTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1380
QY 1594 CGTCAATTTATCAAAAATTAATCTATTTTAAAGGAAAAGCAATATATAAATGATCCAGTGGT 1653
DB 1381 CGTCAATTTATCAAAAATTAATCTATTTTAAAGGAAAAGCAATATATAAATGATCCAGTGGT 1440
QY 1654 GGGTGGCTGCTCAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAAGT 1713

QY	601	GGAAATGTAATCTCGAAGAGAGAACTATTAAAGAGAAAGAAACGCAATTTGAAACAGTCGGA	660	1681	TGTCCTATCAAGAGGAGATAGCAATTACCACTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
DB	601		660	1681		1740
QY	661	ATTGCTCTTACGATTATCACCAGAGAGTGGAAATTTCTGTTTCAAGCCGCTAGTGGGA	720	1741	GGATCTAATATCCAAGTTGATGAAAGTCAGAAGGCTGGTATATTTTGAAGGACCAACAAGAC	1800
DB	661		720	1741		1800
QY	721	ATTTATCACGTAATAAGATGGAGGGCCACAAGGATTTAGCGCAACACCTTTAAAGCCCAAT	780	1801	TCCCTTTTAGAGCATCACCTGTACGTAGTACGTATACGTAAATCTCTGGAGAGGTGACAAGG	1860
DB	721		780	1801		1860
QY	781	CTAGTGGAAACTAGTTGTCCCAACATACGGAATGGATCCAAATTTATGCCCCGCTGATCCA	840	1861	CTGACTCACCGTCTACTCACATTTCTGCTGCATCAGTCAGCAGCTGTGACTTCTTTTATA	1920
DB	781		840	1861		1920
QY	841	GACTGGATTCCTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCCCTTTACAAGCTATCAAGTCCT	1980
DB	841		900	1921		1980
QY	901	GAAAGGAGACTCACTTATGTGCAATGAGCTAGCAACATGGGAAGAGATGCCAGATCA	960	1981	GAAGATGACCCAACTTTGCAAAAACAAAGGAAATTTTGGGCCACCATTTTGGATTCAGCAGGT	2040
DB	901		960	1981		2040
QY	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGGAAGATTTGATAGATATTTCTGGCTATTTGGTGG	1020	2041	CTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
DB	961		1020	2041		2100
QY	1021	TGTCAAAAGCTGAACAACCTCCAGTGTGTGTAATTTCTAGATATTTCTAGATTTCTATATGAAGAA	1080	2101	TTGTATCGGATGCTCTACAGCCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG	2160
DB	1021		1080	2101		2160
QY	1081	AATGATGAATCTGAGTGGAAATTTATTCATGTTACATCCCTATGTTGGAAACAGGAGG	1140	2161	CTGTTTCATATATGGTGTCTCTCAGGTGTCAGTTGGTGAATTAATTCGGTTTAAAGGAGTCAAG	2220
DB	1081		1140	2161		2220
QY	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGATG	1200	2184	CTGTTTCATATATGGTGTCTCTCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA	2280
DB	1141		1200	2184		2280
QY	1201	TCAGAAATATGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGAACTAAT	1260	2281	GGATCTCTGTCACCGAGGCTTAAATTTGAAGGGCCCTTTTAAATATAAAATGGGTCAAATA	2340
DB	1201		1260	2281		2340
QY	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCT	1320	2341	GAATTCAGCATCAGGTGGAGGAGCTCCAAATATCTAGCTTTCTCGATATGATTTTATTGAC	2400
DB	1261		1320	2341		2400
QY	1321	GAGGAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG	1380	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGSCATTA	2460
DB	1321		1380	2401		2460
QY	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTC	1440	2461	ATGCAGAGGTGAGATATCTTCAGGTTGCTATTTGCTGGGGCCCGCCAGTCACCTGTGGATC	2520
DB	1381		1440	2461		2520
QY	1441	ATTGAGTCAGTGCCTGATTTCTGTGACGCCATTAATTTATCTATGAAGAAACAAAGACATC	1500	2521	TTCTATGATACAGGATACCGAAACGTTATATGCGGTACCTTGACAGAGTGAACAGGGC	2580
DB	1441		1500	2521		2580
QY	1501	TGGATAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCAGGAAGGAAATTTGAG	1560	2581	TATTAATTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTA	2640
DB	1501		1560	2581		2640
QY	1561	TTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGTCATTATATACAAAATTTACATCTATT	1620	2641	CTGCTCTTACATGTTTCTCTGGATGAGAAATGTCATTTTGGACATACCAAGTATATCTG	2700
DB	1561		1620	2641		2700
QY	1621	TTAAAGGAAGCAAAATATAAACGATCCAGTGTGGGCTGCTGCTCCAAAGTGAATTCAG	1680	2701	AGTTTATTTAGTGGGGCTGGAAAGCCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2760
DB	1621		1680	2701		2760
QY	1681	TTAAAGGAAGCAAAATATAAACGATCCAGTGTGGGCTGCTGCTCCAAAGTGAATTCAG	1740	2761	AGCATAAGAGTTCTGTAATCGGAGAGAACATTTATGAACTGCATCTTTTGTGCACTACCTTCAA	2820
DB	1681		1740	2761		2820

Db	2461	AGCATAAGAGTTCCTGAATCGGAGAACATTATGAACATGCTCTTTTGCACTACCTTCAA	2520
Qy	2821	GAATAACCTTGGATCACGCTATTGCTCTCTAAAAAGTGATATAATTTTGACCTGTGTAAC	2880
Db	2521	GAATAACCTTGGATCACGCTATTGCTCTCTAAAAAGTGATATAATTTTGACCTGTGTAAC	2580
Qy	2881	TCTCTGGTATACACCTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA	2940
Db	2581	TCTCTGGTATACACCTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA	2640
Qy	2941	ATTGATCATCATTTTGGATACCTGCGCATGTAAACATCTACTCCTGAAAAATAAATGTGGTG	3000
Db	2641	ATTGATCATCATTTTGGATACCTGCGCATGTAAACATCTACTCCTGAAAAATAAATGTGGTG	2700
Qy	3001	CCATGCGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAA	3060
Db	2701	CCATGCGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAA	2760
Qy	3061	TCAATGTACATATTCTGTGAGAGACCCGCAATACCATAAGAAATTACTAAAAA	3120
Db	2761	TCAATGTACATATTCTGTGAGAGACCCGCAATACCATAAGAAATTACTAAAAA	2820

Search completed: April 15, 2006, 03:22:29
Job time : 2427 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3105.2	99.5	3143	14	US-11-151-601-19	Sequence 19, Appl
2	2645.8	84.8	2649	14	US-11-151-601-21	Sequence 21, Appl
3	118	3.8	1346	6	US-09-925-065A-669313	Sequence 669313,
4	113.6	3.6	609	6	US-09-925-065A-743558	Sequence 743558,
C	113.6	3.6	609	6	US-09-925-065A-743559	Sequence 743559,
	94.8	3.0	2338	11	US-11-079-463-1186	Sequence 1186, Ap
7	79.4	2.5	3332	11	US-11-208-288-1	Sequence 1, Appl
8	79.4	2.5	3407	9	US-10-501-035-34	Sequence 34, Appl
9	77.8	2.5	2217	11	US-11-208-288-3	Sequence 3, Appl
10	77.8	2.5	2301	8	US-10-522-789-1	Sequence 1, Appl
C	73.4	2.4	612	6	US-09-925-065A-818446	Sequence 818446,
12	71.6	2.3	4852	14	US-11-136-527-2130	Sequence 2130, Ap
13	71	2.3	2283	11	US-11-208-288-5	Sequence 5, Appl
14	68.6	2.2	1884	11	US-11-079-463-2536	Sequence 2536, Ap
15	65.8	2.1	2814	11	US-11-245-147-168	Sequence 168, Appl
16	65.8	2.1	2814	14	US-11-186-284-54	Sequence 54, Appl
17	63.2	2.0	535	11	US-11-226-869-428	Sequence 428, Appl
18	52.4	1.7	2778	9	US-10-932-182A-5649	Sequence 5649, Ap

```

; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
US-11-151-601-19

Query Match          99.5%; Score 3105.2; DB 14; Length 3143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 11 GCCTCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTCTTAGTGGCGCGCTGCGCGC 70
Db 26 GCCTCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTCTTAGTGGCGCGCTGCGCGC 85
Qy 71 CTGGGTGTGTCACCGCGCGCGCGCGCGAGGAAGCCACTGCAACAGGACCGGAGTGAGGC 130
Db 86 CTGGGTGTGTCACCGCGCGCGCGCGCGAGGAAGCCACTFACAAACAGGACCGGAGTGAGGC 145
Qy 131 GCGCGAGCATGAAGCGCGCGAGGCGCGCTCCATAGCGCACGTCGGGACGCTCGGCGGG 190
Db 146 GCGCGAGCATGAAGCGCGCGAGGCGCGCTCCATAGCGCACGTCGGGACGCTCGGCGGG 205
Qy 191 GCGCGGGGAAGGAATGCAACATGCGACGCGCAATGGAACAGAACAGCTGGGTGTG 250
Db 206 GCGCGGGGAAGGAATGCAACATGCGACGCGCAATGGAACAGAACAGCTGGGTGTG 265
Qy 251 AGATATTTGAACTCGGACCTGTGAGGAGCAATTTGAATCAGAGATCGGCTTAATTGG 310
Db 266 AGATATTTGAACTCGGACCTGTGAGGAGCAATTTGAATCAGAGATCGGCTTAATTGG 325
Qy 311 AGCCCTTTTATGTTAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTGCGGATPACA 370
Db 326 AGCCCTTTTATGTTAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTGCGGATPACA 385
Qy 371 GAAATATATCATGGCTACATGATGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGA 430
Db 386 GAAATATATCATGGCTACATGATGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGA 445
Qy 431 ATGATCCAGATGGACCTCATTCAGACAGAACTTATACCTTGGCATGTCTGGTGAGAA 490
Db 446 ATGATCCAGATGGACCTCATTCAGACAGAACTTATACCTTGGCATGTCTGGTGAGAA 505
Qy 491 GAGAAATACACTGTTTATTTCTGAAATTTCCAAACTATCAATAGAGCAGCAGTCTTAA 550
Db 506 GAGAAATACACTGTTTATTTCTGAAATTTCCAAACTATCAATAGAGCAGCAGTCTTAA 565
Qy 551 TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATT 610
Db 566 TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATT 625
Qy 611 CTCGAGAGAGAACTATTAAGAGAAAGAAAGCAATGGAAACAGTCGGNAATGCTTCTT 670
Db 626 CTCGAGAGAGAACTATTAAGAGAAAGAAAGCAATGGAAACAGTCGGNAATGCTTCTT 685
Qy 671 AGCATATACCAAGGAAGTGGAATCTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGC 730
Db 686 AGCATATACCAAGGAAGTGGAATCTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGC 745
Qy 731 TAAAGATGGAGGCCACAAAGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAA 790
Db 746 TAAAGATGGAGGCCACAAAGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAA 805
Qy 791 CTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGTATCCAGACTGGATTG 850
Db 806 CTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGTATCCAGACTGGATTG 865
Qy 851 CTTTATATACATAGCAACATATTTGGATATCTAAACATCGTAAACAGAGAAAGAGAGAC 910
Db 866 CTTTATATACATAGCAACATATTTGGATATCTAAACATCGTAAACAGAGAAAGAGAGAC 925
Qy 911 TCACTTATGTGCAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCG 970

```

```

Db 926 TCACTTATGTGCAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCG 985
Qy 971 CTACCTTTGTTCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAG 1030
Db 986 CTACCTTTGTTCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAG 1045
Qy 1031 CTGAAACAACTCCAGTGGTGTAAATCTTAGAATCTATATGAAGAAATATGATGAAT 1090
Db 1046 CTGAAACAACTCCAGTGGTGTAAATCTTAGAATCTATATGAAGAAATATGATGAAT 1105
Qy 1091 CTGAGTGGAAATTTATCATGTTACATCCCTATTTGGAAACAAAGAGGCGCAGATTCT 1150
Db 1106 CTGAGTGGAAATTTATCATGTTACATCCCTATTTGGAAACAAAGAGGCGCAGATTCT 1165
Qy 1151 TCCGTTATCTCTAAAAACAGTACAGCAAACTCTTAAAGTCACCTTTTAAAGATGTCA 1210
Db 1166 TCCGTTATCTCTAAAAACAGTACAGCAAACTCTTAAAGTCACCTTTTAAAGATGTCA 1225
Qy 1211 TGATTGATGCTGAAGGAAGGATCATAGATGTATAGATAAGAACTAAATTCACCTTTT 1270
Db 1226 TGATTGATGCTGAAGGAAGGATCATAGATGTATAGATAAGAACTAAATTCACCTTTT 1285
Qy 1271 AGATTCTATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTTGAGGAAA 1330
Db 1286 AGATTCTATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTTGAGGAAA 1345
Qy 1331 ATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAC 1390
Db 1346 ATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAC 1405
Qy 1391 CTGAATTTATTTCCAGTAGAAGATGATGTTATGGAAGGACAGACCTCATTTGAGTCAG 1450
Db 1406 CTGAATTTATTTCCAGTAGAAGATGATGTTATGGAAGGACAGACCTCATTTGAGTCAG 1465
Qy 1451 TGCTCTGAATCTGTGACGCCATTAATTTATGTAAGAAACAAACAGACATCTGGATAAATA 1510
Db 1466 TGCTCTGAATCTGTGACGCCATTAATTTATGTAAGAAACAAACAGACATCTGGATAAATA 1525
Qy 1511 TCCATGACATCTTTTCATGTTTTCCTCAAGTTCACGAGAGAAATTTGAGTTATTTT 1570
Db 1526 TCCATGACATCTTTTCATGTTTTCCTCAAGTTCACGAGAGAAATTTGAGTTATTTT 1585
Qy 1571 CCTCTGAATGCAAAACAGGTTTCCGTCAATTTATACAAATTTACATCTATTTTAAAGAAA 1630
Db 1586 CCTCTGAATGCAAAACAGGTTTCCGTCAATTTATACAAATTTACATCTATTTTAAAGAAA 1645
Qy 1631 GCAATATATAACGATCCAGTGGGCTGCTGCTCCAAAGTGATTTCAAAGTGTCTATCA 1690
Db 1646 GCAATATATAACGATCCAGTGGGCTGCTGCTCCAAAGTGATTTCAAAGTGTCTATCA 1705
Qy 1691 AAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCGGCGCATGATCTAATA 1750
Db 1706 AAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCGGCGCATGATCTAATA 1765
Qy 1751 TCCAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1810
Db 1766 TCCAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1825
Qy 1811 AGCATCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1870
Db 1826 AGCATCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1885
Qy 1871 GTGGCTACTCAATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1930
Db 1886 GTGGCTACTCAATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1945
Qy 1931 GTAAACAGGAAGATCCACATCTGTGCTCCCTTTTAAAGCTATCAAGTCTCTGAAAGATGACC 1990
Db 1946 GTAAACAGGAAGATCCACATCTGTGCTCCCTTTTAAAGCTATCAAGTCTCTGAAAGATGACC 2005
Qy 1991 CAACCTTGGCAAAACAAAGGAATTTTGGGCGCACATTTTGGATTTTCAGCAGGCTCTCTTCTG 2050

```

Db 2006 CAACCTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGCTCTCTCTCCTG 2065
QY 2051 ACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGA 2110
Db 2066 ACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGA 2125
QY 2111 TGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTGTTCAAT 2170
Db 2126 TGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTGTTCAAT 2185
QY 2171 ATGGTGGTCTCAGGTGAGTTGGTGAATTAATCGTTTAAAGAGTCAAGATATTTCCGT 2230
Db 2186 ATGGTGGTCTCAGGTGAGTTGGTGAATTAATCGTTTAAAGAGTCAAGATATTTCCGT 2245
QY 2231 TGAATACCTAGCTCTCTAGTTATCTGTGTAGTGATAGACACAGGATCTGTGTC 2290
Db 2246 TGAATACCTAGCTCTCTAGTTATCTGTGTAGTGATAGACACAGGATCTGTGTC 2305
QY 2291 ACCGAGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAAATAGAAATGAGC 2350
Db 2306 ACCGAGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAAATAGAAATGAGC 2365
QY 2351 ATCAGGTGGAAGGACTCAATATCTAGCTTCTGATGATATTTCAATGATCTAGATCGTG 2410
Db 2366 ATCAGGTGGAAGGACTCAATATCTAGCTTCTGATGATATTTCAATGATCTAGATCGTG 2425
QY 2411 TGGGATCCAGCGGTGCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCGAGGT 2470
Db 2426 TGGGATCCAGCGGTGCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCGAGGT 2485
QY 2471 CAGATATCTCAGGGTCTCTATGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2530
Db 2486 CAGATATCTCAGGGTCTCTATGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2545
QY 2531 CAGGATACACGGAAGCTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATCTTAG 2590
Db 2546 CAGGATACACGGAAGCTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATCTTAG 2605
QY 2591 GATCTGTGGCCATGAAGAGAAATTTCCCTCTGAAACCAATCGTTTACTCTCTTAC 2650
Db 2606 GATCTGTGGCCATGAAGAGAAATTTCCCTCTGAAACCAATCGTTTACTCTCTTAC 2665
QY 2651 ATGGTTCTCTGATGAGATGTCATTTTGCACATACAGTATATTAATCTAGTTTCTTAG 2710
Db 2666 ATGGTTCTCTGATGAGATGTCATTTTGCACATACAGTATATTAATCTAGTTTCTTAG 2725
QY 2711 TGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAG 2770
Db 2726 TGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAG 2785
QY 2771 TTCTGAAATCGGGAGAACATTAATGAATCTTTTGGCACTACCTTCAAGAAACCTTG 2830
Db 2786 TTCTGAAATCGGGAGAACATTAATGAATCTTTTGGCACTACCTTCAAGAAACCTTG 2845
QY 2831 GATCAGTATTTCTCTCTTAAAGTGATATATTTTGAACCTGTGTGAACCTCTCTGTAT 2890
Db 2846 GATCAGTATTTCTCTCTTAAAGTGATATATTTTGAACCTGTGTGAACCTCTCTGTAT 2905
QY 2891 ACACCTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATC 2950
Db 2906 ACACCTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATC 2965
QY 2951 ACATTTTGTATCTGCTCATTAACATCTACTCTGAAATTAATATGTTGGTGCATGAGGG 3010
Db 2966 ACATTTTGTATCTGCTCATTAACATCTACTCTGAAATTAATATGTTGGTGCATGAGGG 3025
QY 3011 GTCTACGTTTGTGTAGTAATCTAATACCTTTAACCCCAATGCTCAAAATCAAAATGATA 3070
Db 3026 GTCTACGTTTGTGTAGTAATCTAATACCTTTAACCCCAATGCTCAAAATCAAAATGATA 3085
QY 3071 CATATTTCTGAGAGCCAGCAATACCAATGAATTTACTTAAAAAATAA 3120
Db 3086 CATATTTCTGAGAGCCAGCAATACCAATGAATTTACTTAAAAAATAA 3135

RESULT 2

US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: MPI00-054PIRCP10MNDIVM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-151-601-21

Query Match 84.8%; Score 2645.8; DB 14; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 214 ATGCGACGACCAATGAAACAGACAGCAGTGGGTGGATATTTGAACTGCGGACTGT 273
Db 1 ATGCGACGACCAATGAAACAGACAGCAGTGGGTGGATATTTGAACTGCGGACTGT 60
QY 274 GAGGAGAAATTTCAATCAGAGGATCGCTTAAATTTGAGCCCTTTTATGTTGAGCGGAT 333
Db 61 GAGGAGAAATTTCAATCAGAGGATCGCTTAAATTTGAGCCCTTTTATGTTGAGCGGAT 120
QY 334 TCCTGGAGTCAGCTTAAAGCTTGGCGGATACAGAAATATCATGGGTACATGATG 393
Db 121 TCCTGGAGTCAGCTTAAAGCTTGGCGGATACAGAAATATCATGGGTACATGATG 180
QY 394 GCTAAGGACCAACATGATTTTCATGTTGTGAAGAGGATGATCCAGATGGACCTCATCA 453
Db 181 GCTAAGGACCAACATGATTTTCATGTTGTGAAGAGGATGATCCAGATGGACCTCATCA 240
QY 454 GACAGATCTTATACCTTGCCTGCTGTGTGAGAACAGAGAAATATCACTGTTTATTTCT 513
Db 241 GACAGATCTTATACCTTGCCTGCTGTGTGAGAACAGAGAAATATCACTGTTTATTTCT 300
QY 514 GAAATTTCCCAAACTATCAATAGCAGCAGTCTTAAATGCTCTCTTTGGAGGCTCTTTTG 573

Db	301	GAAATCCCAAAATATCAATAGAGCAGCTCTTAATGCTCTCTTGGAGGCTCTTTTG	360	Db	1381	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAACAAATATATAACGATCCAGTGGT	1440
Qy	574	GATCTTTTTCAGGCAACACTGAGCTATGGAATGATTTCTCGAGAAGAACTATTTAAGA	633	Qy	1654	GGGCTCGCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT	1713
Db	361	GATCTTTTTCAGGCAACACTGAGCTATGGAATGATTTCTCGAGAAGAACTATTTAAGA	420	Db	1441	GGGCTCGCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT	1500
Qy	634	GAAGAAGAACGATTTGGAACAGTGGGAATTCCTTACAGTATATCAACCAAGAACTGGA	693	Qy	1714	GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAAG	1773
Db	421	GAAGAAGAACGATTTGGAACAGTGGGAATTCCTTACAGTATATCAACCAAGAACTGGA	480	Db	1501	GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAAG	1560
Qy	694	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGGA	753	Qy	1774	CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACCTGATCTAGT	1833
Db	481	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGGA	540	Db	1561	CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACCTGATCTAGT	1620
Qy	754	TTTACGCAACAACCTTTAAGGCCAATCTAGTGGAAACTAGTTGTGCCAACAATACGGATG	813	Qy	1834	TAGTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1893
Db	541	TTTACGCAACAACCTTTAAGGCCAATCTAGTGGAAACTAGTTGTGCCAACAATACGGATG	600	Db	1621	TAGTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1680
Qy	814	GATCCAAAATATATGCCCGCTGATCCAGACTGGATTGCTTTTATATACATGCAACGATATT	873	Qy	1894	ATCAGTCAGCATCTGTGACTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
Db	601	GATCCAAAATATATGCCCGCTGATCCAGACTGGATTGCTTTTATATACATGCAACGATATT	660	Db	1681	ATCAGTCAGCATCTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1740
Qy	874	TGGATATCTAACATCGTTAACAGAGAAAGAGAGCTCACTTATGTGCACAAATGAGCTA	933	Qy	1954	GTGTCCCTTTTACAAGCTATCAAGTCTCGAAGATGACCCAACTTGCACAAAACAAAGGAATTT	2013
Db	661	TGGATATCTAACATCGTTAACAGAGAAAGAGAGCTCACTTATGTGCACAAATGAGCTA	720	Db	1741	GTGTCCCTTTTACAAGCTATCAAGTCTCGAAGATGACCCAACTTGCACAAAACAAAGGAATTT	1800
Qy	934	GCCAAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA	993	Qy	2014	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAAATTTTC	2073
Db	721	GCCAAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA	780	Db	1801	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAAATTTTC	1860
Qy	994	TTTGTAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGT	1053	Qy	2074	TCCTTTGAAAGTACTACTGGGATTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTA	2133
Db	781	TTTGTAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGT	840	Db	1861	TCCTTTGAAAGTACTACTGGGATTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTA	1920
Qy	1054	AAAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT	1113	Qy	2134	CAGCTCGAAAAGAAATATCTCTACTGTGCTGTTTCAATATATGTTGGTCTCTCAGGTGCGATTG	2193
Db	841	AAAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT	900	Db	1921	CAGCTCGAAAAGAAATATCTCTACTGTGCTGTTTCAATATATGTTGGTCTCTCAGGTGCGATTG	1980
Qy	1114	ACATCCCTATGTTGGAACAGAGGGCAGATTCATTCGGTTATCCTAAACAGGTACA	1173	Qy	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATATACCTTAGCCTCTCTAGGT	2253
Db	901	ACATCCCTATGTTGGAACAGAGGGCAGATTCATTCGGTTATCCTAAACAGGTACA	960	Db	1981	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATATACCTTAGCCTCTCTAGGT	2040
Qy	1174	GCAATCTTAAAGTCATTTTAAAGTGTCAAGAAATATGATGATGCTGAGGAAGGATC	1233	Qy	2254	TATGGTTGTAGTATAGACAAACAGGGGATCTGTCCCGAGGGCTTAAATTTGAAGGC	2313
Db	961	GCAATCTTAAAGTCATTTTAAAGTGTCAAGAAATATGATGATGCTGAGGAAGGATC	1020	Db	2041	TATGGTTGTAGTATAGACAAACAGGGGATCTGTCCCGAGGGCTTAAATTTGAAGGC	2100
Qy	1234	ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAGATTCATTTGAAAGGAGTTGAA	1293	Qy	2314	GCCTTTAAATATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATAT	2373
Db	1021	ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAGATTCATTTGAAAGGAGTTGAA	1080	Db	2101	GCCTTTAAATATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATAT	2160
Qy	1294	TATATGCGCAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1353	Qy	2374	CTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT	2433
Db	1081	TATATGCGCAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1140	Db	2161	CTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT	2220
Qy	1354	CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA	1413	Qy	2434	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTTCAGGGTTGCTATT	2493
Db	1141	CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA	1200	Db	2221	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTTCAGGGTTGCTATT	2280
Qy	1414	GATGATGTTATGGAAGGCAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGACGGCACTA	1473	Qy	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTATATG	2553
Db	1201	GATGATGTTATGGAAGGCAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGACGGCACTA	1260	Db	2281	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTATATG	2340
Qy	1474	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533	Qy	2554	GGTCAACCTCGACAGAAAGAAACAGGGCTATTACTTAGGATCTGTGGGCATTCGCAAGCAGAA	2613
Db	1261	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1320	Db	2341	GGTCAACCTCGACAGAAAGAAACAGGGCTATTACTTAGGATCTGTGGGCATTCGCAAGCAGAA	2400
Qy	1534	CCCCAAGTCAGAGAGGAAATTTGAGTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1593	Qy	2614	AGGTTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2673
Db	1321	CCCCAAGTCAGAGAGGAAATTTGAGTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1380	Db	2401	AGGTTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2460
Qy	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1653	Qy	2674	CATTTTGCATACACAGTATATTACTGAGTTTTTTTACTGAGGGCTGGAAGCCATATGAT	2733
				Db	2461	CATTTTGCATACACAGTATATTACTGAGTTTTTTTACTGAGGGCTGGAAGCCATATGAT	2520

Qy	2734	TTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCCGGGAGACATTAT	2793
Db	2521	TTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCCGGGAGACATTAT	2580
Qy	2794	GAACGTGCATCTTTTGGCATTACCTTCAAGAAAAACCTTGGATCACGTATTGCTGCTCTAAAA	2853
Db	2581	GAACGTGCATCTTTTGGCATTACCTTCAAGAAAAACCTTGGATCACGTATTGCTGCTCTAAAA	2640
Qy	2854	GTGATATAA	2862
Db	2641	GTGATATAA	2649

RESULT 3

```

US-09-925-065A-669313/c
; Sequence 669313, Application US/09925065A
; Publication NO. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669313
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-669313

```

RESULT 4

```

RES001 4
US-09-925-065A-743558/c
; Sequence 743558, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30

```

9 T.M.S.B.

RESULTS 6
US-11-079-463-1186
; Sequence 1186, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton

9 T.J.S.B

RESULTS 6
US-11-079-463-1186
; Sequence 1186, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton


```

; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-501-035-34

Query Match      2.5%; Score 79.4; DB 9; Length 3407;
Best Local Similarity 47.7%; Pred. No. 1.3e-10;
Matches 296; Conservative 0; Mismatches 321; Indels 4; Gaps 2;

QY 2270 TAGACAACAGGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAA 2329
DB 1808 TTGATGCGCAGAGGAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGAC 1867
QY 2330 TGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATG 2389
DB 1868 TGGGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAT---TTCAAAATGG 1924
QY 2390 ATTTCATTGACTTTAGATCTGTGGGCATCCAGCGCTGTCTATGGAGGATACCTCTCCC 2449
DB 1925 GATTTGTGACAACAACCAATTTGCAATTTGGGGCTGTCTATGGAGGTCAGTAACCT 1984
QY 2450 TGATGCAATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTTATTCGAGGGTCCAGTCA 2509
DB 1985 CAATGGTCTGGGATCGGGAAGTGGCGTGTTCAGAGTGGAAATAGCGGTGGCGCTGTAT 2044
QY 2510 CTCTGTGGATCTCTATGATACAGGATACAGGAACGTTATATGGTCAACCTGACCCAG 2569
DB 2045 CCGGTGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAG 2104
QY 2570 ATGACAGGGCTATTACTTATAGGATCTGTGGCCATCGAGCAGAGAAAGTTCCCTCTGAAC 2629
DB 2105 AAGACACCTTGACCAATTACAGAAATTCACAGTCATGACAGAGCTGAAAAATTTAAAC 2164
QY 2630 CAAATCGTTTACTGCTTTACATGGTTTCTCGATGAGAATGTCCATTTTGCACATACCA 2689
DB 2165 AAGTTGAGTACCTCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTTCAGCAGTCAG 2224
QY 2690 GTATATPACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCCTC 2749
DB 2225 CTCAGATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATACG 2284
QY 2750 AGGAGACACACAGATAAGANGTCTCGATCGGAGAACATTATGAACGATCTTTTTCG 2809
DB 2285 ATGAAGACCATGGAATAGCTAGCAGCAGCAGCACAACATATATATACCCACATGAGCC 2344
QY 2810 ACTTACCTTCAAGAAACCTTGGATCAGGATACGATTTGCTGCTCTAAAAGTGATATAAATTTGAC 2869
DB 2345 ACTTCATAAACAATGTTTCTCTTTACCT-TAGCAGCTCAAAATACCATGCCATTTAAAG 2403
QY 2870 CTGTGTAGAACTCTCTGGTAT 2890
DB 2404 CTTATTAAACTCATTTTGT 2424

RESULT 9
US-11-208-288-3
; Sequence 3, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: TANIABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2301
```

```

; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-208-288-3

Query Match      2.5%; Score 77.8; DB 11; Length 2217;
Best Local Similarity 48.7%; Pred. No. 2.6e-10;
Matches 242; Conservative 0; Mismatches 252; Indels 3; Gaps 1;

QY 2270 TAGACAACAGGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAA 2329
DB 1649 TTGATGCGCAGAGGAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGAC 1708
QY 2330 TGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATG 2389
DB 1709 TGGGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACA---CAATTTTCAAAAATGG 1765
QY 2390 ATTTCATTGACTTTAGATCTGTGGGCATCCAGCGCTGTCTATGGAGGATACCTCTCCC 2449
DB 1766 GATTTGTGGAACAACAACCAATTTGCAATTTGGGGCTGTCTATGGAGGTCAGTAACCT 1825
QY 2450 TGATGCAATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTTATTCGAGGGTCCAGTCA 2509
DB 1826 CAATGGTCTGGGATCGGGAAGTGGCGTGTTCAGAGTGGAAATAGCGGTGGCGCTGTAT 1885
QY 2510 CTCTGTGGATCTCTATGATACAGGATACAGGAACGTTATATGGTCAACCTGACCCAG 2569
DB 1886 CCGGTGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAG 1945
QY 2570 ATGACAGGGCTATTACTTATAGGATCTGTGGCCATCGAGCAGAGAAAGTTCCCTCTGAAC 2629
DB 1946 AAGACACCTTGACCAATTACAGAAATTCACAGTCATGACAGAGCTGAAAAATTTAAAC 2005
QY 2630 CAAATCGTTTACTGCTTTACATGGTTTCTCGATGAGAATGTCCATTTTGCACATACCA 2689
DB 2006 AAGTTGAGTACCTCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTTCAGCAGTCAG 2065
QY 2690 GTATATPACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCCTC 2749
DB 2066 CTCAGATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATACG 2125
QY 2750 AGGAGACACACAGCATA 2766
DB 2126 ATGAAGACCATGGAATA 2142

RESULT 10
US-10-522-789-1
; Sequence 1, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANIABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2301
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2301)
; OTHER INFORMATION:
US-10-522-789-1

Query Match      2.5%; Score 77.8; DB 8; Length 2301;
Best Local Similarity 48.7%; Pred. No. 2.7e-10;
Matches 242; Conservative 0; Mismatches 252; Indels 3; Gaps 1;

QY 2270 TAGACAAACAGGGGATCCGTGCACGAGGCTTAAATTTGAAGCGCTTTAAATATAAAA 2329
DB 1733 TTGATGCGCAGAGGAAGTGTTACCAAGAGATAAGATCATGCAATCAACAGAGAC 1792
QY 2330 TGGGTCAATAGAAATTCACATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATG 2389
DB 1793 TGGGAACATTTGAGTTGAGATCAATTTGAGCAGCCAGA---CAATTTTCAAAAATGG 1849
QY 2390 ATTCAATGACTTAGTGTGGGCATCCAGGCTGCTCATGAGGATACCTTCTCC 2449
DB 1850 GATTTGTGGACAAACGAATTCGAATTTGGGGCTGTCTATGAGGAGGTACGTAACCT 1909
QY 2450 TGATGCGATTAATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCACTCA 2509
DB 1910 CAATGTCCTGGGATCGGGAAGTGGCGTGTCAAGTGTGGAATAGCCGTGGCGCCTGTAT 1969
QY 2510 CTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTGCACCCCTGAC 2569
DB 1970 CCGGTGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAG 2029
QY 2570 ATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGATTCCCTCTGAA 2629
DB 2030 AAGACAACCTTGACCATACAGAAATTCACAGTCATGACAGAGCTGAAATTTTAAAC 2089
QY 2630 CAAATCGTTACTGCTCTTACATGTTTCTCGATGAGATGTCCTATTTGCACATACCA 2689
DB 2090 AAGTTGAGTACCTCTTATTCATGGAACAGCAGATGATAGCTTCACTTTCAGAGTCAG 2149
QY 2690 GTATATCTAGTGTGTTTGTAGGCTGGAAGCCATATGATTTACAGATCTATCTTC 2749
DB 2150 CTCAGATCTCCAAAGCCCTGTGCGATGTTGGAGTGGATTTCCAGGCAATGTGTTACTG 2209
QY 2750 AGGAGAGACACAGCAT 2766
DB 2210 ATGAAGACCATGGAATA 2226

RESULT 11
US-09-925-065A-818446/c
; Sequence 818446, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 818446

; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2301)
; OTHER INFORMATION:
US-09-925-065A-818446

Query Match      2.4%; Score 73.4; DB 6; Length 612;
Best Local Similarity 98.7%; Pred. No. 1.9e-09;
Matches 74; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2667 GAATTCATTTGCACATACCATATATATCTAGTGTGAGTTTATAGTGAGGCTGGAAGCC 2726
DB 612 GAATTCATTTGCACATACCATATATATCTAGTGTGAGTTTATAGTGAGGCTGGAAGCC 553
QY 2727 ATATGATTTACAGAT 2741
DB 552 ATATGATTTACAGGT 538

RESULT 12
US-11-136-527-2130
; Sequence 2130, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041006 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR FILING DATE: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2130
; LENGTH: 4852
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; ORGANISM: Rattus norvegicus
US-11-136-527-2130

Query Match      2.3%; Score 71.6; DB 14; Length 4852;
Best Local Similarity 53.3%; Pred. No. 2.3e-08;
Matches 162; Conservative 5; Mismatches 134; Indels 3; Gaps 1;

QY 2258 TGGTTCTAGTATGACACACAGGGGATCCTGTCCACCGAGGCTTTAAATTTGAAAGGCGCCT 2317
DB 1812 TRGTAGCTAGCTTTGATGGCAGAKAGAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAA 1871
QY 2318 TTAATATATAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAG 2377
DB 1872 TCACAAAGAGACTTGGACACACTGGAAGTTGAAGATCAAATTTGAAGCAGCC---AGGCAAT 1928
QY 2378 CTCTCGATATGATTTCACTTGAATTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAG 2437
DB 1929 TTTTAAATATGGATTTTGGACACAGCAGCAGTTGCAATTTGGGGCTGGTCAATATGAG 1988
QY 2438 GATACCTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTCTG 2497
DB 1989 GGTACGTAACCTCAATGGTCTCGGATCGGAGTGGCGTGTTCARTGTGGAATACCG 2048
QY 2498 GGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGCTC 2557
DB 2049 TGGGGCCCGTGTCAACGGTGGGAGTACTATGACTAGTATACACAGAGCGTTTACATGGTC 2108
QY 2558 ACCC 2561
DB 2109 TCCC 2112

RESULT 13
US-11-208-288-5
; Sequence 5, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
```

APPLICANT: CHANG, Chiwen
TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
FILE REFERENCE: 39533-0001
CURRENT APPLICATION NUMBER: US/11/208,288
CURRENT FILING DATE: 2005-08-18
PRIOR APPLICATION NUMBER: US 60/605,013
PRIOR FILING DATE: 2004-08-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 2283
TYPE: DNA
ORGANISM: Mus musculus
US-11-208-288-5

Query Match 2.3%; Score 71; DB 11; Length 2283;
Best Local Similarity 49.9%; Pred. No. 2.1e-08;
Matches 207; Conservative 0; Mismatches 205; Indels 3; Gaps 1;
QY 2270 TAGACAAACAGGGGATCTGTCCACCGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAAAA 2329
DB 1715 TTGACGGCAGAGAGTGGTTACCAAGGAGATAGATCATGCATCAACAGAGAT 1774
QY 2330 TGGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATG 2389
DB 1775 TGGGAACACTGGAAGTTGAAGATCAAAATTTGAAGCAGCC---AGGCAATTTGTAAAAATGG 1831
QY 2390 ATTTCATTGACTTATGATGTTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCC 2449
DB 1832 GATTTGTGATAGCAAGCGAGTTGCAATTTGGGGCTGGTCATATGGAGGGTATGTAACT 1891
QY 2450 TGAATGCAATTAATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 2509
DB 1892 CAATGCTCTGGATCGGAAGTGGCGTTCAGTGGGAGATAGCTGTGGCACTGTGT 1951
QY 2510 CTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACTCGACACAGA 2569
DB 1952 CACGTGGGAGTACTATGACTCAGTGTACACAGAGCGTTACATGGGTCTCCCAATTCAG 2011
QY 2570 ATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAC 2629
DB 2012 AAGACAACCTTGACCAATACAGGAACTCAACAGTCATGACGAGCAGTGAACATTTTAAAC 2071
QY 2630 CAAATCGTTTACTGCTTACATGCTTCTTCTGGATGAGATGTCATTTTGCACA 2684
DB 2072 AAGTTGAGTACCTCTCTTATTCATGAACGGCAGATGATATGTTCACTTTCAGCA 2126

RESULT 14
US-11-079-463-2536
Sequence 2536, Application US/11079463
Publication No. US20060073161A1
GENERAL INFORMATION:
APPLICANT: Gary L. Breton
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: PATH00-03DIV2
CURRENT APPLICATION NUMBER: US/11/079,463
CURRENT FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/128,705
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: US 09/540,209
PRIOR FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 10444
SEQ ID NO 2536
LENGTH: 1884
TYPE: DNA
ORGANISM: B. fragilis
US-11-079-463-2536
Query Match 2.2%; Score 68.6; DB 11; Length 1884;
Best Local Similarity 48.2%; Pred. No. 8.5e-08;
Matches 188; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

Matches 225; Conservative 0; Mismatches 239; Indels 3; Gaps 1;
QY 2119 AAGCCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTCAATATATGGTGT 2178
DB 1186 AAACCGCGCGATTTTCGATCCGAAACAAAGAAATACCGGCTATCGTATATACGTATACGGCGT 1245
QY 2179 CTTCAAGTGCAGTTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTCAATACC 2238
DB 1246 CTTCAAGTGCAGTTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTCAATACC 1305
QY 2239 CTAGCCTCTCTAGTTTATGTTGTTAGTGTAGTACAGAGGATCTCTGTCACCGAGGG 2298
DB 1306 ATGCCCAACAAAGGTTTACATCATGTTTACCGTAGACGACGCTGGAGCAGCAATCGGGA 1365
QY 2299 CTTTAAATTTGAAGGCGCTTTTAAATATAAAATGAGTGGTCAAAATAGAAAATTTGACGATCAGGTG 2358
DB 1366 CTCGATTTTGAAGATGTTTACTTTTCCGCCCAATTCCTGAAAGCTCCCTACGTCGATGGGA 1425
QY 2359 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCTGTCGGCATC 2418
DB 1426 AAGGAAACCGAATTCCTGAAAGCTCCCTACGTCGATGGGA---ACCGTATCGGAGTA 1482
QY 2419 CACGGCTGTCTTATGAGGATACCTCTCCCTGATGGCATTAATCGCAGAGTCAAGATATC 2478
DB 1483 CACGGCTGAGCTTTGGCGGTCAATGACCACTGCGCTTCTCTCGCTATCCCGAGATA 1542
QY 2479 TTCAAGGTTGCTTATGCTCGGGCCCCAGTCACTCTGTGATGATCTTCTATGATACAGGATAC 2538
DB 1543 TTTTAAAGTGGCGTGGCGGGTCCCGTCAATTGACTGGGGTATTACGAAGTGTATGATGAC 1602
QY 2539 ACGAAGCTTATATGAGTCACTCCCTGACCAAGATGAACAGGCTATTA 2585
DB 1603 GGAGAACGCTTATATGATATACCCCGCAAGCAATCCGAAAGGATATAA 1649

RESULT 15

US-11-245-147-168
Sequence 168, Application US/11245147
Publication No. US20060030541A1
GENERAL INFORMATION:
APPLICANT: GARCIA, TERESA
APPLICANT: ROMAN ROMAN, SERGIO
APPLICANT: BARON, ROLAND
APPLICANT: CALL, KATHERINE
APPLICANT: THEILHABER, JOACHIM
APPLICANT: CONNOLLY, TIMOTHY
APPLICANT: JACKSON, AMANDA
APPLICANT: BUSHNELL, STEVEN
APPLICANT: RAWADI, GEORGES
TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE
FILE REFERENCE: 37991-0023
CURRENT APPLICATION NUMBER: US/11/245,147
CURRENT FILING DATE: 2005-10-07
PRIOR APPLICATION NUMBER: PCT/IB02/02211
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/281,400
PRIOR FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 246
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 168
LENGTH: 2814
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Homo sapiens fibroblast activation protein, alpha
OTHER INFORMATION: (FAP), mRNA
US-11-245-147-168

Query Match 2.1%; Score 65.8; DB 11; Length 2814;
Best Local Similarity 46.8%; Pred. No. 6.5e-07;
Matches 242; Conservative 0; Mismatches 272; Indels 3; Gaps 1;
QY 2250 AGGTTATGTGTTGTTAGTATGATACAAACAGGGGATCCTGTCACCGAGGCTTAAATTTGA 2309

Db	1903	AGGATGTCATTCGCTTGGTGGATGTCGAGGACAGCTTCCAGGTGACAACTCCT	1962
Qy	2310	AGGCCCTTTAAATATAAATAGGTCAAATAGAAATTGACGATCAGGTGGAGGACTCCA	2369
Db	1963	CTATCCAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGATTACAGCTGTGAG	2022
Qy	2370	ATATCTAGCTTCTCGATATGATTTTCATTTAGACTTAGATTCGTTGGGCATCCACGGCTGGTC	2429
Db	2023	AAA---ATTTCATAGAAATGGGTTCATTTGATGAAAAAAGAAATAGCCATATGGGGCTGGTC	2079
Qy	2430	CTATGGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTGAGATATCTTCAGGGTTGC	2489
Db	2080	CTATGGAGGATACGTTTCATCACTGGCCCTTGCACTGGAACTGGTCTTTTCAAATGTGG	2139
Qy	2490	TATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTA	2549
Db	2140	TATAGCAGTGGTCCAGTCTCCAGTGGGAATATTACGCGTCTGTCTACACAGAGAGATT	2199
Qy	2550	TATGGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCCAAGC	2609
Db	2200	CATGGTCTCCCAACAAAGGATGATAATCTTGACACTATAAGAAATCAACTGTGATGCG	2259
Qy	2610	AGAAAGTTCCTCTGACCAAAATCGTTTACTGCTTTACATGTTTCTCTGGATGAGNA	2669
Db	2260	AAGCAGAAATTTTCAGAAATGTAGACTATCTTCATCCAGGAACAGCAGATGATAA	2319
Qy	2670	TGTCATTTTGCAATACAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATA	2729
Db	2320	TGTGCACTTTCAAACTCAGCAGATGCTAAAGCTCTGGTTTATGCAAAAGTGATTT	2379
Qy	2730	TGATTTACAGATCTTATCTCAGGAGAGACACAGCATA	2766
Db	2380	CCAGGCAATGTGGTACTCTGACCAAGAACCGGCTTA	2416

Search completed: April 15, 2006, 01:59:35
Job time : 1181 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 14, 2006, 10:51:01 ; Search time 90.7 Seconds
(without alignments)
3022.849 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5552
Sequence: 1 aagtgtctaaagctctcgagg.....agaattactaaaaaaaaa 3120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlbp
-Q=/abss/ABSSWEB_epool/US10825632/runat_14042006_092348_3416/app_query.fasta.1
-DB=A Geneseq -QFWT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10825632@cgm_1_1_348 @runat_14042006_092348_3416 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	84.7	882	4 AAB47187	Aab47187 Human DPP
2	4700	84.7	882	5 ABG61591	Abg61591 Human DPP
3	4700	84.7	882	5 AAG78415	Aag78415 Amino aci
4	4700	84.7	882	5 AAE24170	Aae24170 Human dip
5	4700	84.7	882	5 AAE24170	Aau74749 Human pro
6	4700	84.7	882	5 ADI17086	Adi17086 Human NOV
7	4700	84.7	882	6 ABU07720	Abu07720 Human ser
8	4528.5	81.6	883	5 ADI17085	Adi17085 Murine NO
9	4397.5	79.2	831	6 ABU92026	Abu92026 Human pro

10	4353	78.4	824	6 ABU92030	Abu92030 Human pro
11	4118	74.2	782	5 ABB97361	Abb97361 Novel hum
12	3901	70.3	746	8 ADI16334	Adi16334 Human pro
13	3817	68.8	738	8 ADT04044	Adt04044 Human pro
14	3771	67.9	724	5 ABB97362	Abb97362 Novel hum
15	3607.5	65.0	690	5 ABG61594	Abg61594 Human DPR
16	3513.5	63.3	661	5 ABG61596	Abg61596 Human DPR
17	3504	63.1	658	5 ABG61600	Abg61600 Human DPR
18	3333.5	60.0	632	4 AAB93565	Aab93565 Human pro
19	3236	58.3	613	5 ABG61601	Abg61601 Human DPR
20	3087	55.6	587	5 ADR41398	Adr41398 Human CD-
21	2871	51.7	892	5 ABG61602	Abg61602 Human DPR
22	2871	51.7	892	5 ABG61604	Abg61604 Human DPR
23	2871	51.7	892	5 ABB98134	Abb98134 Human PMM
24	2871	51.7	892	8 ADS10951	Ads10951 Human the
25	2871	51.7	969	5 AAE24168	Aae24168 Human dip
26	2870	51.7	863	5 ABG61592	Abg61592 Human DPP
27	2870	51.7	863	5 ADI17083	Adi17083 Human NOV
28	2862	51.5	863	5 ADI16690	Adi16690 Human NOV
29	2862	51.5	863	5 ADI16688	Adi16688 Human NOV
30	2862	51.5	863	8 ADN42344	Adn42344 Human nov
31	2852.5	51.4	580	5 AAE14337	Aae14337 Human pro
32	2842	51.2	863	8 ADN42342	Adn42342 Human nov
33	2835	51.1	830	5 AAE24171	Aae24171 Human dip
34	2833	51.0	869	5 AAE24169	Aae24169 Alternati
35	2821.5	50.8	879	5 ABG61607	Abg61607 Human DPR
36	2821.5	50.8	879	5 ABG61608	Abg61608 Human DPR
37	2821.5	50.8	879	8 ADS10953	Ads10953 Human the
38	2713.5	48.9	847	5 AAE23875	Aae23875 Murine di
39	2708	48.8	854	8 ADS10952	Ads10952 Human the
40	2672.5	48.1	873	8 ADS10479	Ads10479 Human the
41	2523	45.4	493	7 ADE78977	Ade78977 Human pro
42	2505	45.1	755	8 ADQ67811	Adq67811 Novel hum
43	2470	44.5	516	6 ABU92029	Abu92029 Human pro
44	2422	43.6	465	4 AAB47189	Aab47189 Human DPP
45	2413.5	43.5	737	4 AAM38724	Aam38724 Human pol

ALIGNMENTS

RESULT 1
AAB47187
ID AAB47187 standard; protein; 882 AA.
XX
AC AAB47187;
XX
DT 29-JUN-2001 (first entry)
XX
DE Human DPP8.
XX
KW Human; dipeptidyl aminopeptidase; DPP8; prollyl oligopeptidase;
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Active-site 739 /note= "Forms part of Ser-Asp-His catalytic triad"
FT Active-site 817 /note= "Forms part of Ser-Asp-His catalytic triad"
FT Active-site 849 /note= "Forms part of Ser-Asp-His catalytic triad"
FT Active-site 849 /note= "Forms part of Ser-Asp-His catalytic triad"
XX
WO200119866-A1.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-AU001085.
XX
PR 10-SEP-1999; 99AU-00002762.

PR 18-FEB-2000; 2000AU-00005709.
 XX (UNSY) UNIV SYDNEY.
 XX PA Abbott CA, Gorell MD;
 XX PI WPI; 2001-281520/29.
 XX DR N-PSDB; AAC85694.
 XX

XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
 PT substrates, identifying inhibitors of DPP8 catalytic activity which have
 PT therapeutic uses, and for detecting activated T cells.
 XX

XX Claim 1; Fig 2; 78pp; English.

XX This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has
 CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.
 CC Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase,
 CC because it is capable of hydrolysing the peptide bond C-terminal to Pro
 CC in each of these compounds. DPP8 is homologous with human DPP4. DPP8 is
 CC useful for cleaving a substrate, and for detecting an activated T cell.
 CC which involves measuring the level of DPP8 gene expression in a T cell.
 CC The level of DPP8 expression is detected by detecting the amount of DPP8
 CC RNA in the cell. It is also useful for identifying a molecule capable of
 CC inhibiting DPP8 catalytic activity by DPP8. Molecules identified as
 CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
 CC growth hormone deficiency, lowering glucose levels in non-insulin
 CC dependent diabetes mellitus and other disorders involving glucose
 CC intolerance, enhancing mucosal regeneration and as immunosuppressants
 XX

XX Sequence 882 AA;

Alignment Scores:
 Pred. No.: 0 Length: 882
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 84.7% Indels: 0
 DB: Gaps: 0

US-10-825-632-2 (1-3120) x AAB47187 (1-882)

QY 214 ATGGCAGCAGCAATGNAACAGACAGCTGGTGTGAGATATTTGAACTGGGACTGT 273
 Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 QY 274 GAGGAGATATTCGAATCAGGATCGGCTAAATGGAGCCTTTTATGTTGAGCGGTAT 333
 Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCAAGAAATATCATGCGTACATGATG 393
 Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 QY 394 GCTAAGGCACCATGATTTTCATGTTGTGTAAGAGGAATGATCCAGATGGACCTCATTTCA 453
 Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 QY 454 GACAGATCTATTACCTTCGATGCTGGTGAGAACAGAGAAATACACTGTTTATTTCT 513
 Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 QY 514 GAAATCCCAAACTATCATAGCAGCAGCTTTAATGCTCTCTTGGAGCCTCTTTTG 573
 Db 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 QY 574 GATCTTTTTCAGGCACTGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGA 633
 Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
 QY 634 GAAAGAAACGCATTTGMAACAGTCGGAATGCTTCTTACGATTATCACCAAGGAAGTGA 693
 Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGCGAGCGGCACAAGA 753
 Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 QY 754 TTTACGCAACCACTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCTCCACATACCGATG 813
 Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 QY 814 GATCCAAAATTTATGCTCCCGCTGATCCAGACTGGATGCTTTTATATACATAGCAACGATAT 873
 Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 920
 QY 874 TGGATATCTAATCATCGTAACCAAGAGAAAGAGAGACTCATTATGTGCAATAGCTA 933
 Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 940
 QY 934 GCCACATGGAAGAGATGCCAGATCAGCTGGAGTCCCTACCTTGTCTTCCAAAGAGAA 993
 Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
 QY 994 TTTGATAGATATCTGGCTATTGTTGGTGTCCAAAGCTGAAACCACTCCACGTGGTGGT 1053
 Db 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
 QY 1054 AAAATCTTTAGAAATCTATATGAGAAATGATCAATCTGAGGTGGAAATTTATTCATGTT 1113
 Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspLysLeuValGluIleHisVal 300
 QY 1114 ACATCCCTTATGTTGGAACCAAGAGGCGAGATTCATTCGTTATCTTAAACAGGTACA 1173
 Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 QY 1174 GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGATGCTGAAGAGAGATC 1233
 Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGATCTCTATTTGAAGGAGTTGAA 1293
 Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 QY 1294 TATATTGCCAGAGCTGGATGGAGCTCTCGAGGGAATAATGCTCGTCCATCTCTACTAGAT 1353
 Db 361 TyrIleAlaAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 QY 1354 CGTCCCAAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1413
 Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 QY 1414 GATGATGTTATGGAAGCGAGAGACTCATTTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
 Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 QY 1474 ATTATCTATGAGAAACACACACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
 Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 QY 1534 CCCCAGAGTCACAGAGGAAATGAGTTATTTTGGCTCTGATGTCGAAACAGAGTTTC 1593
 Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 QY 1594 CGTCATTTTATACAAAATTCATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGT 1653
 Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 QY 1654 GGGTGGCTCTCTCAAGTGTATTTCAAGTCTCTATCAAGAGAGAGATAGCAATTTACCGAT 1713
 Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
 QY 1714 GGTCAATGGGAAGTTCTTGGCCCGCATGATCTTAATATCCAAAGTTGATGAAGTCAGAGG 1773
 Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520


```
QY 1774 CTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACGTGACGTAGTCACT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValSer 540
QY 1834 TAGTAAATCCCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACATTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAAGATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTACAGCTATCAAGTCTCAAGATGACCACTTGCAAAAACAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGGCACCAATTTTGATTCAGCAGGTCTCTTCTGACTATATCTCTCAGAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTGTATGGATGCTCTACAAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCCTGGAAGAATAATCCTACTGTGCTGTTCATATATATGCTGCTCCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTCAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACCTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGCTGTGTAGTATACAAACAGGGGATCCTGTCAACCGAGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAAGACTCCCAAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCGATATGATTTCAATTCAGTCTAGTATCGTGGGCATCCACGCTGGTCTTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGCTAATTAATGCAGAGTCCAGATATCTTCAGGTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTTCTATGATACAGATACACGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCAACCTGACCAAGATGAAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCTGCAACCAATCGTTTACTGCTCTTACATGCTTCTCGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACCACTATATTACTAGTTTTTTAGTGAGGCTGGAAAGCCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGACACAGCATPAAGATTCCTGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATCGATCTTTTGGCACTACCTTCAGAAAACCTTGATCAGCTATTGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
```

```
Db 881 ValIle 882
RESULT 2
ABG61591
ID ABG61591 standard; protein; 882 AA.
XX
AC ABG61591;
XX
DT 12-AUG-2002 (first entry)
XX
DE Human DPPIV related serine protease DPPP-1.
XX
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX
OS Homo sapiens.
XX
PN WO200231134-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031874.
XX
PR 12-OCT-2000; 2000US-0240117P.
XX
PA (FERR ) FERRING BV.
XX
PI Qi S, Akineanya KO, Riviere PJ, Junien J;
XX
DR WPI; 2002-444178/47.
XX
DR N-PSDB; ABK83322.
XX
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
PS Claim 17; Fig 1; 113pp; English.
XX
CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders.
CC -ABG61591 represent human DPPP proteins
XX
SQ Sequence 882 AA;
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 5 Caps: 0
US-10-825-632-2 (1-3120) x ABG61591 (1-882)
QY 214 ATGGCAGCAGCAATGGAAACAGACGCTGGGTGTGAGATATTTGAACCTCGGACTGT 273
|||||
```

Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Qy	274	GAGGAGAAATTGAATCAACAGATCGGCTAAATTGGAGCCTTTTATGTTGAGCGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Qy	334	TCCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACCCAGAAAATATCATGGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Qy	394	GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Qy	454	GACAGAAATCTATTACCTTGGCATGTCTGTCGAGAACAGAGAAAATACACTGTTTATTCT	513
Db	81	AspArgIleTyr-Tyr-LeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Qy	514	GAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGGAGGCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Qy	574	GATCTTTTTCAGGCAACACTGACTATGGAATGTATCTCGAGAAGAAGCACTATTAA	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
Qy	634	GAAGAAGAACGATTTGGAACAGTCGGAATTCCTCTTACGATTTATCACCAAGAGAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Qy	694	ACATTTCTCTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGAGGGCCACAA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTACGCCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCAACATAC	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Qy	874	TGATATCTAAACATCCTAACACAGAGAAGAGAGACTCATTATGTGCAACATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCCAAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACTTTGTCTCCAAAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Qy	994	TTTGTAGATATTCTGGCTATTGGGTGTCCTCAAGCTGAAACAACTCCACGTGGTGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrTrpSerGlyGly	280
Qy	1054	AAAATTTCTAGAAATCTATACAGAAAATGATGAACTGAGTGGGAAATTTATCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluAsnAspGluSerGluValGluIleIleHisVal	300
Qy	1114	ACATCCCTTATGTTGAAACAAAGAGGGGAGATTCATTCCGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATATGATGCTGAAGGAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGGATTTCTATTGGAAGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCCAGAGCTGGATGAGCTCTGAGGGGAAAATATGCTTGGTCCATCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
Qy	1354	CGCTCCACAGCTCGCTACAGATAGTTGATCTCACCTGAATATTATTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	GATGATGTTATGGAAGCAGAGACTCATTGAGTCAGTCGCTGATTTCTGTGAGCCACACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGCATCTTCTGTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCCAAGTCACGAAGAGAAATTGAGTTTATTTTGGCTCTGTAATCCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAACCATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GGCTCGCTCGCTCCAAAGTGAATTCAGTGTCTCTATCAAGAGAGAGATAGCAATTACCAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Qy	1714	GGTGAATGGGAAGTTCTTGGCGCATCGATCTAATATCCAAGTTGATGAAGTCAGAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	1774	CTGGTATATTTGNAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
Qy	1834	TAGTAAATCTCGAGAGGTGACAGCTGACTGACTCAGCTGCTACTCACATTTCTTGCTGC	1893
Db	541	TyrValAsnProGlyLysValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	1894	ATCAGTCAGACTGTGACTTCTTTATAAGTAAGTATAGTACACAGAGAATCCACTGT	1953
Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTCTCCCTTTACAGCTATCAAGTCCTCAAGATGACCCCACTTCCAAAACAAAGAAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACCATTTTGGATTCAGCAGGTCTCTCTCTGCTACTATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
Qy	2074	TCITTTGAAAGTACTACTGCTGCTGTTTCATATATATGTTGGTCTCTCAGTGCAGTTG	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTGGAAAGAAATATCTACTGCTGCTGTTTCATATATGTTGGTCTCTCAGTGCAGTTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	2194	GTGNAATACGGTTTAAAGAGTCAAGTATTTCCGCTTTGAATACCTAGCCTCTCTAGGT	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Qy	2254	TATGTGCTGTAGTAGACACACAGGGGATCTCTCACCCGAGGCTTTAAATTTTGAAGGC	2313
Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	2314	GCCTTTAAATATATAATGGTCAATAGAAATGACATCAGGTGGAGGACTCCAATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleLeuAspGlnValGluGlyLeuGlnTyr	720
Qy	2374	CTAGCTTCTCGATATGATTTTCATGCTAGATCTGTGGGATCCAGGCTGCTCTAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740

QY	2434	GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAAGGTTCATT	2493	CC	Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760	CC	anthyroid, dermatological, antipsoriatic, antisclerotic, antischmatic,
QY	2494	GCTGGGGCCCACTCTCTGGATCTTCTATGATACAGATACACGGACGTATATG	2553	CC	ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780	CC	anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,
QY	2554	GGTCACCTCACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGACAGAGAA	2613	CC	antiatherosclerotic, anorectic and metabolic in its action. Uses include
Db	781	GlyHisProAspGlnHisGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800	CC	gene therapy, expression or activity of 21953 protein modulator. It is
QY	2614	AAGTCCCTCTGACCAATCGTTTACTGCTCTTACATGGTTCTCTGGATGAGAATGTC	2673	CC	useful for identifying a compound which binds to it and can be used in
Db	801	LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820	CC	preventing, treating or detecting a cellular proliferative or
QY	2674	CATTTCGACATACAGTATATTTACTGAGTTTTTACTGAGGCTGGAGCCATATGAT	2733	CC	differentiative disorder. The 21953 molecules can act as novel diagnostic
Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840	CC	targets and therapeutic agents for controlling disorders associated with
QY	2734	TTACAGATCTATCTCAGGAGACACACATGAAGTTTCCTGAATCGGAGACATTAT	2793	CC	the aberrant activity or degradation of peptide hormones e.g., disorders
Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860	CC	associated with cell differentiation and proliferation such as cancer.
QY	2794	GAACTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTATTGCTCTAAAA	2853	CC	immune function, reproductive, neurological and cardiovascular function.
Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880	CC	The 21953 molecules are thus useful for treating and preventing cellular
QY	2854	GTGATA 2859		CC	proliferative and differentiative disorders, haematopoietic neoplastic
Db	881	ValIle 882		CC	disorders, immune disorders such as autoimmune diseases, diabetes
RESULT 3				CC	mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
AAAG78415				CC	neural disorders, demyelinating diseases, vascular disorders and
ID	AAAG78415	standard; protein; 882 AA.		CC	metabolism or pain disorders. This sequence represents the amino acid
XX				CC	sequence of 21953 human prolyl oligopeptidase
XX	AAAG78415;			XX	SQ
DT	12-APR-2002	(first entry)			
XX					
DE	Amino acid sequence of 21953 human prolyl oligopeptidase.				
XX					
KW	21953 prolyl oligopeptidase; antibody; proline; endopeptidase; cancer;				
KW	cardiovascular disease; autoimmune disease; atopic allergy;				
KW	neural disorder; vascular disorder; prostate disorder; Cytostatic;				
KW	antidiabetic; antiarthritic; antisclerotic; antiinflammatory;				
KW	diabetes mellitus; arthritis; multiple sclerosis; asthma;				
KW	Grave's disease; neuronal disorder; demyelinating disease.				
XX					
OS	Homo sapiens.				
XX					
PN	WO200179473-A2.				
XX					
PD	25-OCT-2001.				
XX					
PF	11-APR-2001; 2001WO-US040483.				
XX					
PR	18-APR-2000; 2000US-0197508P.				
XX					
PA	(MILL-) MILLENNIUM PHARM INC.				
XX					
PI	Meyers RA, Williamson M;				
XX					
DR	WPI; 2002-034353/04.				
XX	N-PSDB; AAH99934.				
DR					
XX	New polypeptides 21953, member of human prolyl oligopeptidase family,				
PT	useful as diagnostic targets and therapeutic agents for controlling				
PT	cancer, lymphoma and leukemia.				
XX					
PS	Claim 1; Page 102-103; 12ipp; English.				
XX					
CC	This invention relates to an isolated 21953 human prolyl oligopeptidase.				

QY	814	GATCCAAATATATCCCGCTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATATT	873	Db	561	IIeSerGlnHisCysAspPhePheIleSerLysTy-SerAsnGlnLysAsnProHisCys	580
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	QY	1954	GTGTCCTCTTCAAGCTATCAAGTCTTGAAGATGACCAACTTGGCAAAACAAGGAAATTT	2013
QY	874	TGGATATCTACATCGTAACAGAGAGAGAGAGACTCATTATGTGCACAATGAGCTA	933	Db	581	ValSerLeuTyrlsLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrlsValHisAsnGluLeu	240	QY	2014	TGGCCACCATTGTTGGATTTCAGCAGGTCCTCTTCCTGACTATATCTCTCCAGAAATTTTC	2073
QY	934	GCCACATGGAAGAGATGCCAGATCAGCTGGATCGCTACCTTCTTCCCAAGAGAA	993	Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrlsThrProGluIlePhe	620
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	QY	2074	TCTTTTGAAGTACTACTGGATTTCATTGTATGGGATGCTCTTACAGGCTCATGATCTA	2133
QY	994	TTTGATAGATATCTCGGCTATTGGTGTGTCCTCAAGCTGAACTCCCTGGTGTGT	1053	Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrlsGlyMetLeuTyrlsProHisAspLeu	640
Db	261	PheAspArgTyrlsSerGlyTyrlsTrpProCysProLysAlaGluThrThrProSerGlyGly	280	QY	2134	CAGCTCGGAAGAAATATCTCTACTGTGTGTTCATATATATGCTGCTCTCAGGTGCAGTTG	2193
QY	1054	AAAACTCTPAGAAATCTATATGAAGAAATGATGATCTGAGTGGAAATTTATCTGTT	1113	Db	641	GlnProGlyLysLysTyrlsProThrValLeuPheIleTyrlsGlyProGlnValGlnLeu	660
Db	281	LysIleLeuArgIleLeuTyrlsGluGluAsnAspGluSerGluValGluIleIleHisVal	300	QY	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCTTAGCTCTCTAGGT	2253
QY	1114	ACATCCCTATCTTGGAAACAAGAGGCGCAGATTCATCTCCGTTATCTCTAAACAGGTACA	1173	Db	661	ValAsnAsnArgPheLysGlyValLysTyrlsPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrlsProLysThrGlyThr	320	QY	2254	TATGTGTTGTAGTATAGACAAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGGC	2313
QY	1174	GCAATCCCTAAAGTCACCTTTTAAGATGTGAGAAATATGATTCATGCTGAAGAGATGC	1233	Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	QY	2314	GCCTTAAATATAAATCGGTCMAATAGAAATTCACGATCAGGTGGAGGACTCCCAATAT	2373
QY	1234	ATAGATGTCATAGATAAGGAATTAATCAACCTTTTGAGATTCCTATTGGAAGGAGTTGAA	1293	Db	701	AlaPheLysTyrlsMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrls	720
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	QY	2374	CTAGCTTCGATATGATTTTCATTGACTTAGATCGTGTGGCATCCAGGCTGCTCTAT	2433
QY	1294	TATATTGCCAGAGCTGGATGACTCTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1353	Db	721	LeuAlaSerArgTyrlsAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrls	740
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrlsAlaTrpSerIleLeuLeuAsp	380	QY	2434	GGAGGATACCTCTCCCTGATGGCATTAATGACAGGTCAGATATCTTCAGGTTGCTATT	2493
QY	1354	CGCTCCAGACTCCCTACAGATAGTGTGATCTCACTCGAATTTATTTATCCAGTAGAA	1413	Db	741	GlyGlyTyrlsLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	QY	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATG	2553
QY	1414	GATGATGTTATGGAAGGAGAGACTCATGAGTCAGTCAGTCTGATTCGTGAGCCACTA	1473	Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrlsAspThrGlyTyrlsGluArgTyrlsMet	780
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	QY	2554	GGTCACCCCTGACCAGAAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA	2613
QY	1474	ATTATCTATGAAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT	1533	Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrlsLeuGlySerValAlaMetGlnAlaGlu	800
Db	421	IleIleTyrlsGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe	440	QY	2614	AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2673
QY	1534	CCCCAAGTCACGAAGAGAAATGAGTTTATTTTGGCTCTGATGCAAAACAGTTTC	1593	Db	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	QY	2674	CATTTTGCACATACCAAGTATATTTACTAGTGTGTGTAGTGGGCTGGAAAGCCATATGAT	2733
QY	1594	CGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGT	1653	Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrlsAsp	840
Db	461	ArgHisLeuTyrlsIleThrSerIleLeuLysGluSerLysTyrlsLysArgSerSerGly	480	QY	2734	TTACAGATCTATCTCAGGAGAGACACACATACAGTTCTCTGAATCGGAGAACATTTAT	2793
QY	1654	GGGTGCTGCTCTCAAGTATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATACCAGT	1713	Db	841	LeuGlnIleTyrlsProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrls	860
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500	QY	2794	GAATCGATCTTTTGCACTACCTTCAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA	2853
QY	1714	GGTGAAGGGAAGTTCTTGGCCGGCATGATCTAATATCCAAAGTTGATGAAGTCAGAAGG	1773	Db	861	GluLeuHisLeuLeuHisTyrlsLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	QY	2854	GTGATA 2859	
QY	1774	CTGGTATATTTGAGGCCCAAGACTCCCTTTTAGAGCATCACCTGATAGTACGTAGTCAGT	1833	Db	881	Vallile 882	
Db	521	LeuValTyrlsPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrlsValValSer	540	RESULT 4			
QY	1834	TACGTAATCTCGAGAGTGCAAGGCTGACTGACCGTGGCTACTCATTCTTGTGTC	1893	AAE24170			
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrlsSerHisSerCysCys	560	ID AAE24170			
QY	1894	ATCAGTCAGCACTGTGATCTCTTTTAAAGTAAGTATAGTAAACCAGAGAATCCACACTGT	1953	XX AAE24170;			
				XX			

DT 23-SEP-2002 (first entry)
 DE Human dipeptidyl peptidase 8 (DPP8) protein.
 XX Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
 KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
 KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
 KW antiviral; enzyme.
 XX Homo sapiens.
 OS WO200234900-A1.
 PN 02-MAY-2002.
 PD
 XX 29-OCT-2001; 2001WO-AU001388.
 PF 27-OCT-2000; 2000AU-00001078.
 PR (UNSY) UNIV SYDNEY.
 XX Abbott CA, Gorrell MD;
 PI WPI; 2002-454646/48.
 DR N-PSDB; AAD38956.
 XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
 PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
 PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
 PT infection.
 XX Example; Fig 1; 91pp; English.
 PS The present invention relates to dipeptidyl peptidase (DPP) proteins and
 CC polynucleotides encoding such proteins. The DPP peptides are useful for
 CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
 CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
 CC rejection and HIV (human immuno deficiency virus) infection. The present
 CC sequence is human DPP8 protein
 XX
 SQ Sequence 882 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 882 -
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 84.7% Indels: 0
 DB: 5 Gaps: 0
 US-10-825-632-2 (1-3120) x AAE24170 (1-882)
 QY 214 ATGGCAGCAGCAATGGAAACAGACGCTGGGTGTGAGATATTGTAACCTGGCGACTGT 273
 DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
 QY 274 GAGGAGAAATTAATCAGATCGGCCTAAATGGAGCCTTTTATGTTGAGCGGTAT 333
 DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGCTACATGATG 393
 DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaSerThrArgLysTyrHisGlyTyrMetMet 60
 QY 394 GCTAAGGCACCATGATTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453
 DB 61 AlAlaAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 QY 454 GACAGATCTATTACCTTGCATGCTGTGTGAGACAGAGAAATACACTGTTTATTCT 513
 DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 QY 514 GAAATCCCAAAACTATCAATAGACGACGAGTCTTAATGCTCTCTTGGAGCCTCTTTTG 573

Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 QY 574 GATCTTTTTCAGCAACACTGGACTATGGATGTATTCTCGAGAGAAGAACTATTAAAGA 633
 Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
 QY 634 GAAAGAAAACGATTTGGAAACAGTCGGAAATCTCTTACGATTTATCCCAAGGAGCTGGA 693
 Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCACGTAAAGATGAGGGCCACAAAGGA 753
 Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 QY 754 TTTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGTTCTCCACATACGATG 813
 Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAAGATATT 873
 Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 QY 874 TGGATATCTAATCGTAAACAGAGAAAGGAGACTCACTTATGTGCAATGACGCTA 933
 Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 QY 934 GCCAAATCGAAGAAGATGCCAGATCAGCTGAGTGCCTACTTGTTCCTCAAGAGAA 993
 Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
 QY 994 TTTGATAGATATTCGGCTATTGGTGTGTCCTCAAGAGCTGAAACACTCCCGAGTGGT 1053
 Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 QY 1054 AAATTTCTAGATTTCTATATCAAGAAATGATGATCTGAGGTGGAAATTTATCATGTT 1113
 Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 QY 1114 ACATCCCTTATGTTGAAAACAAAGGAGGCGAGATTCATTCCGTTATCTCTTAAACAGGTACA 1173
 Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 QY 1174 GCAATCTTAAATCACTTTTAAGATGTCAGAAATATGATGATGCTGAAGGAGGATC 1233
 Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 QY 1234 ATAGATGTCATAGATAAGGAGTAAATTCACCTTTTGGAGATTTATTTTGAAGGAGTTCGAA 1293
 Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 QY 1294 TATATTCCACAGACTGGATGGACTCCTCGAGGAAATATGCTTGTGCTCATCTACTAGAT 1353
 Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
 Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 QY 1414 GATCATGTTTATGAAAAGGAGAGACTCATTGAGTCAGTGCCTGATTCTGTGACGCCACTA 1473
 Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 QY 1474 ATTATCTATCAAGAAACACAGACATCTGGATAAATATATCCATGACATCTTTTCATGTTT 1533
 Db 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
 QY 1534 CCCAAAGTCACGAAGAGGAAATGAGTTATTTTTCCTCTGAATGCAAAACAGCTTTC 1593
 Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 QY 1594 CGTCATTTATCAAAATTTACATCTATTATTAAGGAAAGCAATATAAAGATCCAGTGTGT 1653

Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGly 480
QY 1654 GGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTCAAGAGAGAGATAGCAATACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleGlyGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAATATCAAGTTGATGAAGTCAGAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCACAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TAGCTAAATCTCGAGAGCTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCCAGCACTGTGACTCTCTTTATAGTAAGTATAGTAAACCAAGAGAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTACAGCTATCAAGTCCCTCAAGATGACCCACTTGCAGAAACAAGCAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCACCATTTTGGATTCAGAGGTCCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTGGAATTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCCTGGAAGAAATATCTCTACTGTCTGTTCATATATATGTTGCTCCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCGGTTTAAAGAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTCATAGACAACAGGGGATCCTGTCCAGCGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAATATAAATGGTCAATAGAAATGACGATCAGGTGGAAGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCAATTCAGTTCGATCGTGGCATCCAGCGTGGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAAGCTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCTGACAGAAATGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAA 2613
Db 781 GlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AGATTCTCCCTCGAACCAGAAATCGTTTACTGTCTTACATGTTTCTCTGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCATACACAGATATATTTACTGAGTTTTTTAGTGGGGCTGGAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

QY 2734 TTACAGATCTATCTCTCAGAGAGACACAGCATAAGAGTCTCTGATCCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAAGTCGATCTTTTGCATCTACCTTCAGAAAACTTCGGATCACATATGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 5
AAU74749
ID AAU74749 standard; protein; 882 AA.
XX AAU74749;
AC
XX
XX
DT 09-APR-2002 (first entry)
XX
DE Human protease PRTS-9 protein sequence.
XX
XX Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
XX Cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
XX inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
XX cell proliferative disorder; developmental disorder; epilepsy;
XX Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
XX reproductive disorder; endometriosis.
OS Homo sapiens.
XX
XX WO200198468-A2.
XX
XX 27-DEC-2001.
XX
XX 13-JUN-2001; 2001WO-US019178.
XX
XX 16-JUN-2000; 2000US-0212336P.
PR 22-JUN-2000; 2000US-0213953P.
PR 29-JUN-2000; 2000US-0215396P.
PR 07-JUL-2000; 2000US-0216821P.
PR 14-JUL-2000; 2000US-0218946P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
XX Deleagne AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan PA;
XX Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
XX Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
XX Kallick DA;
XX WPI; 2002-090437/12.
XX N-PSDB; ABK12892.
XX
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
XX the diagnosis, treatment and prevention of gastrointestinal (e.g.
XX gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
XX (e.g. cancer) disorders.
XX
XX Claim 1; Page 140-142; 17pp; English.
XX
XX The present invention relates to twenty one new human proteases, referred
XX to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
XX invention are useful in the diagnosis, treatment and prevention of
XX gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
XX disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
XX infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
XX syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
XX cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
XX epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
XX disease and reproductive e.g. infertility and endometriosis disorders.
XX Numerous other examples of each disorder are given in the specification.

CC The present protein sequence represents the human protease PRTS-9 protein
CC of the invention
XX
SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
Dbs: 5 Gaps: 0

US-10-825-632-2 (1-3120) x AAU74749 (1-882)

QY	214	ATGCGAGCAGCAATGGAACACAGACAGCTGGGTGTTGAGATATTTGAACTCGGACTGT	273
Db	1	MetAlaAlaAlaMetGluThrGluInLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGAAATATTGAATCAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAGCTCTTGGCCGATACAGAAATATCATCGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACTCATTTCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAACTTATTACCTTGGCATGCTGGTGAGAACAGAGAAATACATGTTTATTCT	513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeupheTyrSer	100
QY	514	GAAATTCCTCAAACTATCAATAGACGACAGCTTATGCTCTTTGGAGCCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTCAGGCAACACTGACTATGGAATGATTCTCGAGAAGAAGAACTATTAAAG	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
QY	634	GAAAGAAACGATTTGGAACAGTCGGAATTCCTTACGATTATCACCAAGAAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTGTCCCAACATACCGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAAATTTATGCCCCCTGATCCGACTGGATTGCTTTTATACATACCAACGATAT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAACATCGTAAACAGAGAAGAGGAGACTCATCTTATGTGCGCATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGAAGAAGATCCGATCAGCTGAGTCGCTACCTTTGTTCTCCAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaIleThrPheValLeuGlnGlu	260
QY	994	TTTGTAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrTrpProSerGlyGly	280
QY	1054	AAAAATCTTAGAATCTTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113

Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAAACAAGAGGGCAGATTCATTCGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAAAATCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGCTGAGGAAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATTAAGGAACATAATCAACCTTTTTCAGATCTTATTTGAAGGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGACTCTCTGAGGAAATATGCTTGGTCCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTyrTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGACTCGCCTACAGATAGTGTGATCTACCTGAAATTTATTTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTGCCTGATCTGTGAGCGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAGTCACGAAGAGGAATTCAGTTTATTTTGCCTCTGAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAATATCATCTATTTTAAAGGAACAAATAAATACCATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGCTCTCCCTCCCAAGTATTTCAAGTGCTCTATCAAGAGAGAGATAGCAATATACCAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAAGTCTTTCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTGTAGTCACT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValIleSer	540
QY	1834	TAGTAAATCCTGGAGAGGTGACAGGCTGACTGACCTGACCTGCTACTCATCTTGTCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACGTGCTCTTCTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGAAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACCAATTTTGGATTTCAGAGGTCTCTCTGCTGACTATACCTCCCGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	2074	TCCTTTTCAAGATCTACTGGATTTACATTTGTTATGGGATGCTCTACAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATATGTTGGTCTCAGGTGAGTTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660

or physiological responses in a cell, tissue, organ or organism. Specifically, it refers to the use of biologically active fragments for diagnostic and prognostic assays and furthermore in the treatment of diverse pathological conditions. The present invention describes novel human and murine NOVX proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOVX polypeptides, polynucleotides and antibodies are useful in treating or preventing NOVX-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including cytosstatic, cardiant, antiinflammatory, immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic, antiasthmatic, nephrotropic, antiarthritic, hepatotropic, neuroprotective, nontropic, antibacterial, virucide, antiparasitic, relaxant and anticonvulsant. In addition, they are useful in screening assays to identify small molecules that modulate or inhibit, for example, neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. This polypeptide is a homologue of a human NOVX protein of the invention.

Sequence 882 AA;

Alignment Scores:

Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 5 Gaps: 0

US-10-825-632-2 (1-3120) x ADI17086 (1-882)

QY	214	ATGCGACGACCAATGGAACAGACAGCTGGGTGTGAGATATTGGAACCTGCGGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys	20
QY	274	GAGGAGATATTGAATCAGGATCGGCTAAATGGAGCCCTTTTATGTTGACGGTAT	333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAAGCTGCTGCGGATACACAGAAATATCATGGCTACATGATG	393
DB	41	SerTrpSerGlnLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAATCTATTACCTTGGCATGTCTGCTGAGAACAGAGAAATATACATGTTTATTTCT	513
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAATTCCTCCAAAATATCAATAGACGAGCTCTTAATGCTCTCTTGGAACTCTTTTG	573
DB	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGGCMACACTGACTATGGAATGCTATCTCCGAGAGAGAACTATTAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
QY	634	GAAGAAACCGCATTTGAAACAGTCGGAATGCTTCTTACGATTATCACCAAGGAGTGGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAATTTATACGTAAAGATGGAGGGCCACAAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180

QY	754	TTTAGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTGCCAACATACGGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCAAAATTTATGCCCGCTGATCCAGACTCGATTGCTTTTATACATAGCAACGATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTTAACATCTGTAACAGAGAAAGAGACATCATTTATGTCCACAATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgAspLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGGAAGAAGATGCCAGATCAGCTGAGCTGCTTACCTTTGTTCTCCAGAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
QY	994	TTTGATAGATATTCTGGCTATTGCTGCTCCAAAGCTGMAAACAACTCCCAAGTGGTGGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	1113
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCAATCCGTTATCTTAAACAGGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATCTCATAGATAAGGAACATAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTGA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATGCCAGAGCTGGAGTGGACTCCTGAGGGAAATAATGTTGGTTCATCTCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGAGCTCGCTACAGATAGTGTGTGATCTCACCTGGAATTTATTTATCCACAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATCATGTTATGGAAGGACAGACATCATTTGAGTCCAGTCCCTGATTTCTGTGACGCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCACGAGGAGAAATGATTTATTTTGGCTCTGAAATGCAAAACAGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGGCTGGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAAAGAGAGATAGCAATACCATG	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGATGGGAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTGGATGAAGTCAGAGG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGTGTATTTTGAAGSCCAAGAGACTCCCTTTTAGAGCATCACCTGTACCTGTAGTCAGT	1833
DB	521	LeuValTyrPheGluGluThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
QY	1834	TACGTAAATCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCTGC	1893

Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560		XX
Qy	1894	ATCAGTCAGCACTGTGACTCTTTTATTAAGTAAGTAGTAACACAGAGAATCCACACTGT	1953		XX
Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580		XX
Qy	1954	GTGTCCTCCCTTCAAGCTATCAAGTCTGAGATGAGCCCAACTTGCACAAACAAGAAATTT	2013		XX
Db	581	ValSerLeuLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600		XX
Qy	2014	TGGGCACATTTGGATTGAGAGCTCTCTCCCTGCTATCTCCACCAAAATTTTC	2073		XX
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuPronAspPyrThrProGluIlePhe	620		XX
Qy	2074	TCTTTTGAAGFACACTGGATTTACATTGTATGGGATGCTCTACAGGCTCATGATCTA	2133		XX
Db	621	SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640		XX
Qy	2134	CAGCCTGGAAGAAATATCCTACTGCTGCTCATATATATGATGGTCTCAGGTGCAGTTG	2193		XX
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660		XX
Qy	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCAGGT	2253		XX
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680		XX
Qy	2254	TATGTGTTGTAGTATAGACACACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGC	2313		XX
Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700		XX
Qy	2314	GCCTTTAATAATAAGGTCAATAGAAATTCACGATCAGGTGGAAGACTCCCAATAT	2373		XX
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720		XX
Qy	2374	CTAGCTTCGATATGATTCATTGACTTAGATCGTGGGCATCCACGGCTGCTCTAT	2433		XX
Db	721	LeuAlaSerArgTyrAspPheIleAspLysAspArgValGlyIleHisGlyTrpSerTyr	740		XX
Qy	2434	GGAGGATACCTCTCCCTGATGGCATTATGACAGAGTCAGATATCTTCAGGGTTGCTATT	2493		XX
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760		XX
Qy	2494	GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATG	2553		XX
Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780		XX
Qy	2554	GGTCACCCCTGACAGCAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA	2613		XX
Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800		XX
Qy	2614	AAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2673		XX
Db	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820		XX
Qy	2674	CATTTTGCACATACCACTATATTACTAGTATTTTGTAGTGGGCTGGAAAGCCATATGAT	2733		XX
Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValAlaGalaGlyLysProTyrAsp	840		XX
Qy	2734	TTACAGATCTATCCTCAGGAGACACAGCATTAAGATTCTGTAATCGGAGAACATTAT	2793		XX
Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860		XX
Qy	2794	GAAGTCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCAAAA	2853		XX
Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880		XX
Qy	2854	GTGATA	2859		XX
Db	881	ValIle	882		XX
RESULT 7				ABU07720	XX

ABU07720 standard; protein; 882 AA.	ID
ABU07720;	XX
19-MAY-2003 (first entry)	AC
Human serine protease HIPHUM46.	XX
Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;	XX
serine protease activity modulation; dipeptidyl peptidase activity;	XX
musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;	XX
Alzheimer's disease; paraneuronal palsy; Huntington's disease;	XX
amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;	XX
irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;	XX
haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;	XX
colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;	XX
multiple sclerosis.	XX
Homo sapiens.	OS
Key	XX
Location/Qualifiers	XX
259..260	PH
Region	FT
Note= "Paired glutamates of the beta propeller domain"	FT
739	FT
/label= Catalytic_serine_residue	FT
817	FT
/label= Catalytic_aspartate_residue	FT
849	FT
/label= Catalytic_histidine_residue	FT
GB2374869-A.	XX
30-OCT-2002.	XX
22-JAN-2002; 2002GB-00001404.	XX
23-JAN-2001; 2001GB-00001760.	XX
(GLAX) GLAXO GROUP LTD.	XX
Edbrooke MR, Lewis AP;	XX
WPI; 2003-150703/15.	XX
N-PSDB; ABX12255.	XX
Identifying modulators of serine protease activity useful for treating	XX
musculoskeletal diseases, by contacting cell expressing a novel serine	XX
protease polypeptide with a compound and monitoring serine protease	XX
activity.	XX
Claim 10; Page 26-29; 38pp; English.	PS
The invention relates to a method of identifying a substance that	XX
modulates serine protease activity, comprising contacting a cell such as	CC
a neuronal cell, lung cell, intestinal cell or a cell infected with a	CC
virus, expressing a serine protease polypeptide (HIPHUM 46), or its	CC
variant having dipeptidyl peptidase activity, or a serine protease	CC
isolated from the cell with a test substance, and monitoring for serine	CC
protease activity. The method is useful for identifying a substance that	CC
modulates serine protease activity. A modulator of the serine protease is	CC
useful in the manufacture of a medicament for treatment of Pophylaxis of	CC
a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus	CC
infection, Alzheimer's disease, paraneuronal palsy, myotonic	CC
dystrophy, Huntington's disease or amyotrophic lateral sclerosis.	CC
Additional disease that may be treated using modulators of the serine	CC
protease include malabsorption syndromes, irritable bowel syndrome, lung	CC
disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,	CC
rectal polyps, small bowel tumours, colorectal tumours, anaemia,	CC
dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple	CC
sclerosis. The present sequence represents the amino acid sequence of the	CC
human serine protease HIPHUM46	XX
Sequence 882 AA;	XX

Alignment Scores:		0	Length: 882
Pred. No.:	4700.00	882	
Score:	100.0%	0	
Percent Similarity:	100.0%	0	
Best Local Similarity:	100.0%	0	
Query Match:	84.7%	0	
DB:	6	0	
US-10-825-632-2 (1-3120) x ABU07720 (1-882)			
QY	214	ATGGCAGCAGCAATATGGAACACAGACCTGGGTGTTGAGATATTTGAACTCCGCACTGT	273
DB	1	MetAlaAlaLeuMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys	20
QY	274	GAGGAGAAATATGGAATCACAGATCGCCCTAAATATGGAGCCCTTTATGTTGAGCGGTAT	333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAGCTGCTTGCAGATACAGAAATATCATCGCTACATGATG	393
DB	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysIleHisGlyTyrMetMet	60
QY	394	GCTAAGGCACCAATGATTTTCATCTTGTGAAGAGGAATGATCCAGATGGACCTCATCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAATCTATTACCTTGCATCTCTGGTGAGACAGAGAAATACATGTTTATCT	513
DB	81	AspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAAATCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTG	573
DB	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGCAACACTGCACTATGGAATGATTTCTCGAGAAGAAGCACTATTAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
QY	634	GAAAGAAACCGATTTGGAACAGTCGGAATTCCTTTCATGATTTATCACCAGGAAGTGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGTTAGTGGATTTATCACCTAAAGATGGAGGCCACAGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCCAACATACGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATATGCCCCCGTATCCAGACTGGATTCGTTTATACATAGCAACGATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAACTGTAAACAGACAGAAAGGAGACTCCTTATGTGCAACATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
QY	994	TTTGTATGATATTTGGCTATTGGTGTGTCCTCAAGCTGAGTGGTGTGCTT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAATTTCTAGAAATCTATATCAAGAAATGATGATCTGAGGTGGAATTTATTCATGTT	1113
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGAAACAAAGGAGGCGAGATTTCCTCCGTATCTCAAAACAGGTACA	1173

DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAAAATCTAAAGTCACATTTTAAGATGTCAGAAATATGATTCATGCTGAAGAAAGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGAACTAATTCACCTTTTCAGATTTCTATTGGAAGAGTTGAA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGACCTCTGAGGAAAAATATGCTGGTCCATCTCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCGACCTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGGAAGCGAGACTCATTCAGTCAGTCGCTGATTTCTGTGAGCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAGTCACGAAGAGAGAAATTCAGTTTATTTTTCCTCTGAATCCAAAACAGGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACCATCCAGTGGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGCTGCTCTCCCAAGTCATTTCAAGTGCTCTCAAGAGGAGATAGCAATATACCAGT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGACGTAGTCAGT	1833
DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	1834	TACGTAATTCCTGGAGAGTGACAAGGCTGACCTGCTGCTGCTACTCATCTTTCTGTCG	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGACTTCTTTTATAGTAAAGTATAGTAAACCAAGAAATCCACTGT	1953
DB	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTT	2013
DB	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACCATTTTGGATTCAGCAGGTCTCTCTCTGACTATACCTCTCCAGAAATTTTC	2073
DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
QY	2074	TCTTTTCAAAAGTACTCTGGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTA	2133
DB	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCCTGGAAAGAAATATCTCTGCTGCTGTTTCATATATATGCTGCTCCTCAGTGGT	2193
DB	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	2194	GTGATTAATCGTTTAAAGGAGTCAAGTATTTCCGCTTCAATACCTCAGCTCTCTAGGT	2253
DB	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680

human and murine NOVX proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOVX polypeptides, polynucleotides and antibodies are useful in treating or preventing NOVX-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including cytostatic, cardiac antiinflammatory, immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic, antiasthmatic, nephrotropic, antiarthritic, hepatotropic, neuroprotective, nootropic, antibacterial, virucide, antiparasitic, relaxant and anticonvulsant. In addition, they are useful in screening assays to identify small molecules that modulate or inhibit, for example, neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. This polypeptide is a homologue of a human NOVX protein of the invention.

US-10-825-632-2 (1-3120) X ADI17085 (1-883)

Qy	214	ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGGAGATATTTGAAACTCCGACACTGT	273
Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaGluCys	20
Qy	274	GAGGAG---AATATTGAATCACAGATCGGCCTAAATTGGAGCGCTTTTATGTCAGACGG	330
Db	21	GluGluGlyAsnGlyGluSerGlnAspArgProLysLeuGluProPheTyrValGluArg	40
Qy	331	TATTCTCGAGTCAGCTTAAAAAGCTGTTCCCGATACCCAGAAAAATATCATGGCTACATG	390
Db	41	TyrSerTrpSerGlnLeuLysLysLeuAlaaspThrArgLysTyrHisGlyTyrMet	60
Qy	391	ATGGCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCAT	450
Db	61	MetAlaLysAlaProHisAspPheMetPheValLysArgThrAspProAspArgProHis	80
Qy	451	TCAGACAGAAATCTATTACCTTCCTGCCATGCTCGTCGAGAACAGAGAAAATACACTGTTTAT	510
Db	81	SerAspArgValTyrTyrLeuAlaMetSerGlyGluasnArgGluasnThrLeuPheTyr	100
Qy	511	TCGTGAATTCCTCCAAAACATCAATAGACAGCAGCTCTTAATGCTCTCTCTGGAGCCTCTT	570
Db	101	SerGluIleProLysThrIleasnArgAlaAlaValLeuMetLeuSerTrpLysProLeu	120
Qy	571	TTGGATCTTTTTCAGCAACACTGCATATGGAATGTATCTCGAGAGAGAACTATTA	630
Db	121	LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeu	140
Qy	631	AGAGAAAGAAACGCATTTGGACAGTCGGAATTCCTTACGATATATCACCAAGGAAGT	690
Db	141	ArgGluArgLysArgIleGlyThrValGlyIleAlaLysArgAspTyrHisProGlySer	160
Qy	691	GGAAACATTTCTGTTTCAACCCGTAGTGGAAATTTTATCACGTAAAGAATGGAGGGCCACA	750
Db	161	GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisIleLysAspGlyGlyProHis	180
Qy	751	GGATTTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCCACATACCG	810
Db	181	GlyPheThrGlnGlnProLeuArgProasnLeuValGluThrSerCysPAsnIleArg	200

Db 561 CysLeuSerArgHisCysAspPhePheLeuSerLysTyrSerAsnGlnLysAsnProHis 580
QY 1951 TGTGTGTCCTTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAGGA 2010
Db 581 CysValSerLeuTyrLysLeuSerSerProGluAspAspProValHisLysThrLysGlu 600
QY 2011 TTTTGGGCACCATTTTGATTCACGAGGTCCTCTCTGACATPATACTCCTCAGAAATT 2070
Db 601 PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIle 620
QY 2071 TTTCTCTTTTGAAGTACTGATTTACATTGTATGGGATGCTCTACAGGCTCATGAT 2130
Db 621 PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAsp 640
QY 2131 CTACAGCTGGAAAGAAATATCTACTGTCTGTCTATATATATGTTGCTCTCAGGTGCAG 2190
Db 641 LeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGln 660
QY 2191 TTGGTGAATAATCGTTTAAAGGAGTCAAGTATTTCGCTTGAATACCTAGCCTCTCTA 2250
Db 661 LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeu 680
QY 2251 GGTATGTGTTGTAGTATGATACACAGGGGATCCTGTCCACCGAGGCTTAAATTTGAA 2310
Db 681 GlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu 700
QY 2311 GCGGCTTTAAATATAATTAATGGTCAATATGAATGACGATCAGTGAAGACTCCAA 2370
Db 701 GlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln 720
QY 2371 TATCTAGCTTCCTGATATGATTTTCAATGACTGTAGATCGTGGGCATCCACGGCTGTC 2430
Db 721 TyrLeuAlaSerGlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSer 740
QY 2431 TATGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGTTGCT 2490
Db 741 TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla 760
QY 2491 ATTGTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGTTAT 2550
Db 761 IleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyr 780
QY 2551 ATGGGTCACTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCA 2610
Db 781 MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla 800
QY 2611 GAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAT 2670
Db 801 GluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsn 820
QY 2671 GTCCATTTTGCACATACAGTATATTTACTGAGTCTTTTACTGAGGCTGGAAGCCATAT 2730
Db 821 ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr 840
QY 2731 GATTTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCTGATCGGAGNACT 2790
Db 841 AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis 860
QY 2791 TATGAACGATCTTTTTCATCTACTCCTCAAGAAACCTTGGATCAGATATGCTGCTCTA 2850
Db 861 TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeu 880
QY 2851 AAAAGTGATA 2859
Db 881 LysValIle 883

RESULT 9
ABU92026
ID ABU92026 standard; protein; 831 AA.
XX
AC ABU92026;
XX

DT 15-JUL-2003 (first entry)
DE Human protein modification and maintenance molecule-6 (PMMH-6).
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;
KW antiinflammatory; thyromimetic.
XX
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Rankumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan EM, Lee EA, Griffin JA, Li JX, Yue H;
PI Sprague JW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kabie AB, Yue H;
PI Marquis JP, Yao MG, Richardson RW, Tang YF, Jin P, Chen D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI; 2003-430274/40.
XX
DR N-PSDB; ACA92421.
XX
PT New human protein modification and maintenance molecules (PMMH), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
PS Claim 1; Page 242-243; 311pp; English.
XX
CC The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMH), and the polynucleotide
CC sequences encoding them. A total of 40 PMM polypeptides (designated PMM
CC -1 to PMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMM polypeptides of the invention
XX
SQ Sequence 831 AA;
Alignment Scores:
Pred. No.: 0 Length: 831
Score: 4397.50 Matches: 831
Percent Similarity: 94.2% Conservative: 0
Best Local Similarity: 94.2% Mismatches: 0
Query Match: 79.2% Indels: 51

DB:	6	1	Gaps:	
US-10-825-632-2 (1-3120) x ABU92026 (1-831)				
QY	214	ATGGCAGCAGCAATGAAACAGACAGCTGGGTGGTGGATATTTGAACTCGCGACTGT	273	
Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20	
QY	274	GAGGAGAAATTCATCAATCACAGGATGGCTTAATTTGGAGCCCTTTTATGTGGCGGTAT	333	
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40	
QY	334	TCCTGGAGTCAGCTTAAAGCTGTTCGCCGATACACAGAAATATCATGTCATCATGATG	393	
Db	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60	
QY	394	GCTAAGGCCACCATGATTTTCATGTTTGTGGAGGAGTATCCAGATGGACCTCATTTCA	453	
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80	
QY	454	GACAGATCTATTACCTCTGCATGCTGCTGAGACAGAGAAATACACTGTTTATTTCT	513	
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100	
QY	514	GAATTTCCCAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG	573	
Db	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu	120	
QY	574	GATCTTTTTCAGCAACACTGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAG	633	
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140	
QY	634	GAAGAAAGCGCATTCGACAGTCGGAATGCTCTTACGATTTATCACCAGGAAGTGA	693	
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGAATTTATACGTAAAGATGAGAGGCCCAAGGA	753	
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	
QY	754	TTTACGCAACAACTTTAAGGCCAATCTAGTGGAACTAGTTGTCCCAATACGAGATG	813	
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	
QY	814	GATCCAAAATTTAGCCCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATTT	873	
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	
QY	874	TGGATATCTAACATCTGAACAGAGAAAGGAGACTCACTTATGTGCACAATGAGCTA	933	
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	
QY	934	GCCAACTGGAAGAAGATCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAAAGAGAA	993	
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260	
QY	994	TTTGATAGATTTCTGGCTATTGGTGGTCCCAAAGCTGAAACAACTCCAGTGGTGT	1053	
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	
QY	1054	AAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113	
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	
QY	1114	ACATCCCTATGTTGGAACAAGGAGGCGAGATTCATTCCTCGTTATCTTAAACAGGTACA	1173	
Db	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	
QY	1174	GCAATCCTTAAGTCACTTTTAAAGATGTCAGAAATTAATGATGCTGAGAGGAGATC	1233	
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	
QY	1234	ATAGATGTCATAGTAAGGAACATAATTCAACCTTTTGAGATTCTTATTTGAAGGAGTTGAA	1293	
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	
QY	1294	TATATTGCCAGACTGGATGGACTCTCTGAGGGAATAATATGCTTGCTTACTACTAGAT	1353	
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	
QY	1354	CGCTCCAGACTCGCTACAGATAGTGTGTCTCACCTGAATTTATTTATCCAGTAGAA	1413	
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	
QY	1414	GATGATCTTATGAAAGGAGAGACTCATTTGATGTCAGTGCCTGATTTCTGTCCGCCACTA	1473	
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	
QY	1474	ATTATCTATCAAGAAACACAGACATCTCGATAAATATCCATGACATCTTTTCATGTTTT	1533	
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	
QY	1534	CCCCAAAGTCACGAAGAGAAATTTGATGTTTTTTTGTCTCTGAATGCAAAACAGTTTC	1593	
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	
QY	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1653	
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	
QY	1654	GGGTGCTCTCTCAAGTGTATTTCAAGTGTCTTATCAAGAGGAGATAGCAATACCAGT	1713	
Db	481	GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer	500	
QY	1714	GGTGAATGGGAAGTTCTTGGCCGGCATGATCTAATATCCAAAGTTGATGAAGTGAAGG	1773	
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	
QY	1774	CTGTATATTTTGAAGCCCAAGACTCCCTTTAGAGCATCACCTGATGATGTCAGT	1833	
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540	
QY	1834	TACGTAATCTCGAGAGGTGACAAGCTGACTGCGGTGCTACTCACATTTCTTGCTGC	1893	
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	
QY	1894	ATCAGTCAGCACTGTGACTTTCTTTAAGTAAAGTATAGTAACCAAGAAATCCACATGT	1953	
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	
QY	1954	GTGTCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAGAAATTT	2013	
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	
QY	2014	TGGGCCACCATTTTGGATTCCAGCAGGTCTCTTCTGCTACTATCTCTCCAGAAATTTTC	2073	
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620	
QY	2074	TCCTTTGAAAGTACTACTGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA	2133	
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	
QY	2134	CAGCTCGGAAGAAATATCTCTACTGTGCTGTTTCATATATATGTTGGTGTCTCAGTGCAGTTG	2193	
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660	
QY	2194	GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCTCTCTAGT	2253	
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680	
QY	2254	TATGCTGTGTAGTGTATAGACACAGGGATCTCTGTCCCGAGGGCTTTAAATTTGAAGGC	2313	
Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700	
QY	2314	GCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT	2373	

Db 701 AlaPheLysTyrLysMet----- 706
QY 2374 CTAGCTTCTCGATATGATTTTCATTGACTTAGTCGTGGGCATCCACGGCTGCTCTAT 2433
Db 706 ----- 706
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATCCAGAGGTCAGATATCTTCAGGGTTGCTATT 2493
Db 707 -----ValAlaIle 709
QY 2494 GCTGGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACAGGACGTTATATG 2553
Db 710 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 729
QY 2554 GGTCACTCCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
Db 730 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 749
QY 2614 AAGTTCCCTCTCGAACCAAAATCGTTTACTCTCTTACATGGTTTCTGTGATGAGATGTC 2673
Db 750 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 769
QY 2674 CATTGTCACATACCAGTATATTAAGTGTGAGGCTGGAAGCCATATGAT 2733
Db 770 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 789
QY 2734 TTACAGATCTATCTCTCAGGAGACACAGATAGAGTTCTCTGAATCGGAGAACATTAT 2793
Db 790 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 809
QY 2794 GAACTGCATCTTTTGCATCTACCTTCAAGAAACCTTGGATCAGTATTGCTCTCAAAA 2853
Db 810 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 829
QY 2854 GTGATA 2859
Db 830 ValIle 831

RESULT 10
ID ABU92030
XX ABU92030 standard; protein; 824 AA.
XX AC ABU92030;
XX DT 15-JUL-2003 (first entry)
XX DE Human protein modification and maintenance molecule-10 (PMM-10).
XX KW Human; protein modification and maintenance molecule; PMM; cancer;
XX KW cell proliferation disorder; atherosclerosis; neurological disorder;
XX KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
XX KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
XX KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
XX KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
XX KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
XX KW antiinflammatory; thyromimetic.
XX OS Homo sapiens.
XX PN WO2003031939-A2.
XX DX 17-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032850.
XX PR 12-OCT-2001; 2001US-0329689P.
XX PR 25-OCT-2001; 2001US-0335703P.
XX PR 09-NOV-2001; 2001US-034887P.
XX PR 28-NOV-2001; 2001US-0334345P.
XX PR 06-DEC-2001; 2001US-0337451P.
XX PR 14-DEC-2001; 2001US-0340584P.
XX PA (INCY-) INCYTE GENOMICS INC.

XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hatalla AJA, Chawla NK, Lehr-Nason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI; 2003-430274/40.
DR N-PSDB; ACA92425.
XX New human protein modification and maintenance molecules (PMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX Claim 1; Page 249-251; 31pp; English.
XX The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMM polypeptides (designated PMM
CC for PMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal disorders (e.g. hypothyroidism, Cushing's
CC PMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMM polypeptides of the invention
XX SQ Sequence 824 AA;
Alignment Scores:
Pred. No.: 0 Length: 824
Score: 4353.00 Matches: 824
Percent Similarity: 93.4% Conservative: 0
Best Local Similarity: 93.4% Mismatches: 0
Query Match: 78.4% Indels: 58
DB: 6 Gaps: 1
US-10-825-632-2 (1-3120) x ABU92030 (1-824)
QY 214 ATGCAGCAGCAATGAAACAGACAGCTGGGTGTGAGATATTGAACTGCGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATCAGAGATCGGCTAAATGGAGCCTTTTATGTTGACGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTGCCGATACAGAAATATCATGCGTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTGTGAGAGGAATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCCATGCTGCTGGTGGAGACAGAGAAATACACTGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATCCCAAACTATCAATAGAGCAGAGCTCTTAATGCTCTTGGAGCCCTTTTGG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerIlePheProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTTCGAGAGAGAACTATTAGA 633

121 AspLeuPheGln----- 124
124 ----- 124
694 ACATTTCGTGTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAGATGGAGGCCACAAGGA 753
124 ----- 124
754 TTACGCAACAACCTTTAAGCCCAACTAGTGGAACTAGTTGTCCTCCACATACGGATG 813
125 -----GlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 142
814 GATCCAAATATATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
143 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 162
874 TGGATATCTAACATCGTAAACAGAGAAGAGAGACTCACTTATGTGCACAATGAGCTA 933
163 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 182
934 GCCAATCGGAAGAAGATGCCAGATCGACTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAA 993
183 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 202
994 TTTGATAGATATTCGGCTATTGGTGGTGTCCAAAGCTGAACAACCTCCAGTGGTGGT 1053
203 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 222
1054 AAAATCTTAGAATCTATATAGAGAAATGATCAATCTGAGGTGGAAATTTATTCATGTT 1113
223 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 242
1114 ACATCCCTTATCTTGGAAACAAGGAGGCGAGATTCACTCCGTTATCTCTAAACAGGTACA 1173
243 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 262
1174 GCAATCTCAAAGTCACTTTAAGATGTCAAGATATGATGATGCTGAGGAAGGATC 1233
263 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 282
1234 ATAGATGTCATAGATAAGACTAAATCAACCTTTTGAGATCTCTATTGGAAGAGTTGAA 1293
283 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 302
1294 TATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT 1353
303 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 322
1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAAATATTATTTATCCAGTAGAA 1413
323 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 342
1414 GATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
343 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 362
1474 ATTATCTATGAAGAACACACACATCTGGATAATATCCATGATCTTTCTGATGTTTTT 1533
363 IleIleTyrGluGluThrThrAspIleIleTrpIleAsnIleHisAspIlePheHisValPhe 382
1534 CCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
383 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 402
1594 CGTCATTTATACAAATTCATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGT 1653
403 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 422
1654 GGCCTCCCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713

423 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 442
1714 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1773
443 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 462
1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGT 1833
463 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 482
1834 TACGTAAATCTCGAGAGGTGACAAGGCTGACACCGTCTGCTACTCACATTTCTTGCTGC 1893
483 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 502
1894 ATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGT 1953
503 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 522
1954 GTGTCCCTTTACAGCTATCAAGTCTCGAAGTACCACTTCGACCACTTCGCAAAACAAGAAATTT 2013
523 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 542
2014 TGGGCCACCATTTTGGATTACAGAGTCTCTCTCTGCTACTATATCTCTCCAGAAATTTTC 2073
543 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 562
2074 TCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA 2133
563 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 582
2134 CAGCCTGGAAAGAAATATCTCTGCTGTGTTTATATATATGTTGCTCTCAGGTGCACTTG 2193
583 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 602
2194 GTCAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCCTCTCTAGGT 2253
603 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 622
2254 TATGTGTTGTAGTATAGACACAACAGGGGATCCTCTCACCGAGGCTTTAAATTTTGAAGGC 2313
623 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 642
2314 GCCTTTAAATATAAATGCGTCAATAGAAATTCAGATAGATGAGTGGAGGACTCCATAT 2373
643 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 662
2374 CTAGCTTCTCGATATCATTTTCAATTCATGATCTGTGGGATCCACGGCTGGTCTCTAT 2433
663 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 682
2434 GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGTCAAGATATCTTCAGGGTTGCTATT 2493
683 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 702
2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATG 2553
703 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 722
2554 GGTCAACCTCAGCAGAAATGAAAGGCTATTACTTAGGATCTGTGGCCATCCAGACAA 2613
723 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 742
2614 AAGTTCCTCTCAACCAATCGTTTACTGCTTCTACATGTTTCTCTGATGAGATGTC 2673
743 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 762
2674 CATTTTGCACATACCAAGTATATTACTGAGTGTATTTTAGTGAGGCTCGGAAGCCATAT 2733
763 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 782
2734 TTACAGATCTATCTCAGGAGAGACACAGCAATAGAGTTCCTGAATCGGAGAACATTTAT 2793
783 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 802

QY	2794	GAAC	TGCATCTTTTGCAC	TACCTTCAAGAAAACTTGGATCAGTATTCCTGCTTAAAA	2853
Db	803	GluLeuHisLeuLeuHisLeuLeuHisLeuLeuGlnGluAsnLeuGlySerArgileAlaAlaLeuLys	822		
QY	2854	GTGATA	2859		
Db	823	ValIle	824		
RESULT 11					
ID	ABB97361	standard; protein; 782 AA.			
XX	ABB97361;				
XX	27-JUN-2002	(first entry)			
XX	Novel human protein	SEQ ID NO: 629.			
XX	Human; anti-anemic; vulnerary; anti-inflammatory; immunomodulator;				
XX	neuroprotective; cerebroprotective; cytosstatic; rheumatic; gene therapy;				
XX	expressed sequence tag.				
OS	Homo sapiens.				
XX	WO200222660-A2.				
XX	21-MAR-2002.				
XX	10-SEP-2001; 2001WO-US026015.				
XX	11-SEP-2000; 2000US-00659671.				
XX	(HYSE-) HYSEQ INC.				
XX	Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;				
XX	Xue AJ, Yang Y, Wehrman T, Drmanac RT;				
XX	WPI: 2002-292408/33.				
XX	N-PSDB; ABN32547.				
XX	An isolated polynucleotide for treating diseases associated with its				
XX	encoded polypeptide such as cancer and multiple sclerosis.				
XX	Example 2; SEQ ID NO 629; 509pp; English.				
XX	The present invention provides the protein and coding sequences of 444				
XX	residues of human proteins. These were isolated from expressed sequences tags				
XX	(ESTs). They can be used to stimulate cell growth, to regulate regrowth				
XX	haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth				
XX	e.g. in burn treatment, to regulate the immune system e.g. to treat				
XX	multiple sclerosis, to regulate activin or inhibin e.g. to treat				
XX	infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke				
XX	and cancer, to screen for drugs, to treat inflammatory conditions e.g.				
XX	rheumatoid arthritis, and to treat nervous system disorders e.g.				
XX	Parkinson's disease. The present sequence is a protein of the invention				
XX	Sequence 782 AA;				
Alignment Scores:					
Pred. No.:	0	Length:	782		
Score:	4118.00	Matches:	782		
Percent Similarity:	88.7%	Conservative:	0		
Best Local Similarity:	88.7%	Mismatches:	0		
Query Match:	74.2%	Indels:	100		
DB:	5	Gaps:	1		
US-10-825-632-2 (1-3120) x ABB97361 (1-782)					
QY	214	ATGGCAGCAGCAATCGAACAGACAGCTGGGTGCTGAGATATTTGAACTGGGACTCT	273		
Db	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20		

QY	274	GAGGAGATATTTGAATTCACAGATCGCGCTAAATTTGAGCCCTTTTATGTTGACCGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyValGluArgTyr	40
QY	334	TCTCGAGTCAGCTTTAAAAAGCTGCTTCCGATACCCAGAAAAATATATCATGCTCATGATG	393
Db	41	SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyMetMet	60
QY	394	GCTAAGGCCACCATGATTTTGTGTGTAAGAGGAATGATCCAGATGACCTCATPTCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAAATCTATTACCTTGCCTATGCTGTCGTCGAGAACAGAGAAAAATACACTGTTTATCT	513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAATTTCCCAAACTATCAATAGACGACAGCTCTTAATGCTCTCTTCCGAGCCCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGCAACACTGACATGATGTAATGTATTCGAGAGAGAACTATTAAAGA	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
QY	634	GAAAGAAACCGATTGGACAGCTCGGAATTCCTTACGATATCACCAAGAGGGA	693
Db	141	GluArgLysArgIleLeuThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTCAAGCGGTAGTGAATTTATCAGCTAAAGATGAGGGCCACAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTCTCCCAACATACGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAAATATTATGCCCGCTGATCCAGACTCGATGCTTTTATATACATAGCAACAT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGATATCTAAACATCGTAACACAGAGAGAAAGAGACTCACTTATGTGCAATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGGAAGAAGATGCCAGATCAGCTGAGTCGCTACTTTGTTCTCCAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
QY	994	TTTGATAGATATTCTGGCTATTGGGTGTCCTCAAAAGCTGAAACAACTCCCAGTGGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAATTTTAGAATCTATGACAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
QY	1114	ACATCCCTTATGTTGGAAACAAGAGGGGAGATTCATTCCTGTTTCTTAAACAGAGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAAAATCTTAAAGTCATTTTAAAGATGTCAGAAATATGATGATCTGAAGAGAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGGAACATAATTCAACTTTTGTAGATTCATTGTAAGAGGTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380

QY	1354	CGCTCCAGACTCGCTACAGATAGTGTTCATCTCACTGAATATTATATCCAGTAGAA	1413	Db	658	-----ValAlaIle	660
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	QY	2494	GCTGGGGCCCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTATATG	2553
QY	1414	GATGATGTTATGGAAGCAGAGACTCATTTGAGTCAGTGCCTGATTCCTGAGCCACTA	1473	Db	661	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	680
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	QY	2554	GGTCAACCTGACACAGATGAACAGGGCTATTACTAGATCTGTGGCCATGCCAAGCAGAA	2613
QY	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT	1533	Db	681	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	700
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisaspIlePheHisValPhe	440	QY	2614	AGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTCTTCTGGATGAGAATGTC	2673
QY	1534	CCCCAAGTCAACGAGAGAAATGAGTTTATTTTGGCTCTCGAATGCAAAACAGGTTTC	1593	Db	701	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	720
Db	441	ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysIysThrGlyPhe	460	QY	2674	CATTTTGCATACACAGATATATTACTGAGTTTTTTTAGTGGGTGGAAAGCCATATGAT	2733
QY	1594	CGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAGCAAAATATAACGATCCAGTGT	1653	Db	721	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValAlaGlyLysProTyrAsp	740
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGly	480	QY	2734	TTACAGATCTATCCTCAGGAGACACACAGCATAGAGTTCTCTGAATCGGAGAACATTAT	2793
QY	1654	GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGAGATAGCAATTACCAGT	1713	Db	741	LeuGlnIleTyrProGlnGluHisTyrLeuGlnAsnLeuGlySerArgIleAlaAlaLeuIys	760
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500	QY	2794	GAATCGATCTTTTGCACCTACCTCAAGAAAACCTTGGATCACGTTATTGCTCTATAAA	2853
QY	1714	GGTGAATGGGAAGTTCTTTGGCCGCATCGATCTAATATCCAACTTGATGAAGTCAGAAG	1773	Db	761	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuIys	780
Db	501	GlyGluTrpGluValLeuLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	520	QY	2854	GTGATA	2859
QY	1774	CTGGTATATTTGAAGGCACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833	Db	781	ValIle	782
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540	RESULT 12			
QY	1834	TACGTAATCTCGGAGGTGCACAGCTGACTGACCGTGTCTACTCACATTTCTGCTGC	1893	ID	AD116334	standard; protein; 746 AA.	
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys	560	XX	AD116334		
QY	1894	ATCAGTCAAGCTGTGACTCTTTTATAGTAAAGTATAGTAACCAAGAAATCCACACTGT	1953	AC	AD116334		
Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	XX			
QY	1954	GTGTCCTTTACAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTT	2013	DT			
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	XX			
QY	2014	TGGGCCACATTTTGGATTCAGCAGGTCTCTCTGCTACTATCTCTCCAGAAATTTTC	2073	DE			
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	XX			
QY	2074	TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA	2133	XX			
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	OS			
QY	2134	CAGCCTGGAAGAAATATCTATCTGTGCTGTTCATATATGTTGTCCTCAGGTGCAGTTG	2193	XX	WO2003100016-A2.		
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln-----	657	XX	04-DEC-2003.		
QY	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTTAGGT	2253	XX			
Db	657	-----	657	XX	22-MAY-2003; 2003WO-US016498.		
QY	2254	TATGTGTTGTAGTAGACACAACAGGGGATCTGTCCACGAGGGCTTAAATTTGAAGGC	2313	XX			
Db	657	-----	657	XX	22-MAY-2002; 2002US-0383491P.		
QY	2314	GCCTTTAAATATAAAATGGGTCAAATAGAAATTCACGATCAGGTGGAAGGACTCCAATAT	2373	XX	24-JUN-2002; 2002US-0391378P.		
Db	657	-----	657	XX	22-JUL-2002; 2002US-0397921P.		
QY	2374	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATCCACGGCTGCTCTAT	2433	XX	(INCY-) INCYTE CORP.		
Db	657	-----	657	XX	Khare R, Bulloch SA, Swarnakar A, Elliott VS, Marquis JP, Mason PM, Chawla NK, Rankumar J, Kable AE, Hafalia AJA, Lee SY, Tran UK, Yue H, Becha SD, Griffin JA, Chang H, Jiang X, Jackson AA, Richardson TW, Lal PG, Yao MG, Lu Y, Warren BA, Jin P, Wilson AD, Gietzen KJ,		
QY	2434	GGAGGATACCTCTCCCTGATGCGATTAAATGCAGAGGTTCAGATATCTTCAGGGTGTCTATT	2493	XX	WPI; 2004-035124/03.		
				XX	DR N-PSDB; AD116386.		
				XX	New protein modification and maintenance molecules, useful for diagnosing		

PT or treating e.g. peptic ulcer, hypertension, rheumatic fever, AIDS, or
PT Cushing's syndrome, Alzheimer's disease, multiple sclerosis, stroke
XX cancers.

PS Claim 1: SEQ ID NO 19; 419pp; English.

XX The invention comprises the amino acid and coding sequences of human
CC protein modification and maintenance molecules (PMM). The DNA and
CC protein sequences of the invention are useful for the diagnosis and
CC treatment of disorders associated with expression of PMM, such as:
CC gastrointestinal disorders (e.g. peptic ulcer and Crohn's disease),
CC cardiovascular disorders (e.g. hypertension and congenital heart
CC disease), autoimmune or inflammatory disease (e.g. AIDS and anaemia),
CC developmental disorders (e.g. Cushing's syndrome and tubular acidosis),
CC epithelial disorders (e.g. eczema and scabies), neurological disorders
CC (e.g. Alzheimer's disease and multiple sclerosis), infections and cancer.
CC The present amino acid sequence represents a human PMM protein of the
XX invention.

SQ Sequence 746 AA;

Alignment Scores:
Pred. No.: 0 Length: 746
Score: 2901.00 Matches: 746
Percent Similarity: 84.6% Conservative: 0
Best Local Similarity: 84.6% Mismatches: 0
Query Match: 70.3% Indels: 136
DB: 6 Gaps: 1

US-10-825-632-2 (1-3120) x ADI16334 (1-746)

QY 214 ATGGCAGCAGCATGGAAACAGAACAGCTGGGTGGTGGATATTTGAACTCGGAGCTGT 273
DB 1 MetAlaAlaMetGluThrGluGluLeuGlyValGluLeuPheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATCAGAGATCGGCCTAAATTTGGAGCTTTTATGTGAGCGGTAT 333
DB 21 GluGluAenilleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAGAAATATCATGGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTATTCA 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCCATGCTGTGGTGAGAACAGAGAAATACACTGTTTATTCT 513
DB 81 AspArgLleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAenThrLeuPheTyrSer 100
QY 514 GAAATCCCAAACTATCAATAGACGACGCTTAAATGCTCTTGGAGGCTCTTTTG 573
DB 101 GluileProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGGCAACTCGACTATGGAATGATTTCTCGAAGAAGAACTATTAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGATTTGGAACAGTCGGAATTCCTTTACGATTATCACCAAGAAAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyLeuAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAAAGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTTACGCAACACCTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAenilleArgMet 200
QY 814 GATCCAAATATATGCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATAT 873

DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTTAACATCGTTAACAGAGAGAAAGGAGACTCACTTATGTGCAATGAGCTA 933
DB 221 TrpIleSerAsnilleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATCGGAAGAAGATGCCAGATCAGCTGAGTCGCTACTCTTGTTCCTCAAGAAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTTAGAAATCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTTATCTTGGAAACAAAGAGGAGGAGATTCATTCCGTTATCTTAAACACAGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACATTTTAAAGATGTCAAAATATATGATTGATGCTGAAGGAAGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACATAATTCACCTTTTTCAGATTTCTATTGAGGAGTTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGGACTCTCGAGGAAATATGCTGCTCCATCCCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTCACAGATAGTGTGATCTCACCTGAAATTTATTTCCCAAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTTGAAGAAGCAGAGACTCATTCAGTCAGTCGCTGATCTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCAAAGTCACGAAGAGGAATTCAGTTTATTTTGGCTCTGATTCGAAACACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATCAAAATATACATCTATTTTAAAGGAAGCAAAATATAAATCCATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAGAGTATTCAGTGCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
DB 481 GlyLeuProLysProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY 1714 GGTGAATCGGAAGTTCCTGGCCGCGATCGATCTAATATCCAGTGTGATGAGTCAGAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTGAAGGCACAAAGACTCCCTTTTAGAGCATCATCCCTGTACGTAGTCAGT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValIser 540
QY 1834 TACGTAATCTCGGAGAGGTGACAAAGCTGACCGCTGCTACTCACAATTTCTGTCTGC 1893
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACTGTGACTTCTTTTAAAGTAAGTATAGTAAACCAAGAGAAATCCCACTGT 1953
DB 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580

QY 1954 GTGTCCTTTTCAAGCTATCAAGTCTCAAGATGACCCAACTTGCAGAAACAAAGGAATTT 2013
DB 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCACATTTTGGATTACAGAGGTCCTCTCCAGTACTATATCTCTCAGAAATTTTC 2073
DB 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTACATTTATGGAATGCTCTACAGGCTCATGATCTA 2133
DB 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATATCTACTGTCTGTTCATATATGTTGCTCTCAGGTGCGAGTTG 2193
DB 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGTTGAAATACCCCTAGCCTCTCAGGT 2253
DB 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTATAGACACACAGGGATCCTGTCCAGGGCTTAAATTTGAAGC 2313
DB 681 TyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTTAAATATAAATCGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATAT 2373
DB 701 AlaPheLysTyrLysMet----- 706
QY 2374 CTAGCTTCTCGATATGATTTCACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT 2433
DB 706 ----- 706
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATT 2493
DB 706 ----- 706
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATG 2553
DB 706 ----- 706
QY 2554 GGTCAACCTTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGAGAA 2613
DB 706 ----- 706
QY 2614 AAGTTCCCTCTGAACCAAAATCGTTTACTGTCTTACATGCTTCTCGGATGAGAATGTC 2673
DB 706 ----- 706
QY 2674 CATTTTGCATACACAGTATATTAATCTAGATTTTGTAGGGCTGGAAGCCATATGAT 2733
DB 706 ----- 706
QY 2734 TTACAGATCTATCTCAGGAGAGACACAGATAGAGTTCTGTAATCGGAGAACATTTAT 2793
DB 707 -----IleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 724
QY 2794 GAAGTGCATCTTTTGCACTTACCTTCAAGAAACCTTGGATCACGATATTGCTGCTCAAA 2853
DB 725 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 744
QY 2854 GTGATA 2859
DB 745 ValIle 746
RESULT 13
ADT04044
ID ADT04044 standard; protein; 738 AA.
XX AC
XX ADT04044;
DT 30-DEC-2004 (first entry)
XX

DE Human protein modification and maintenance molecule protein SeqID24.
XX protein modification and maintenance molecule; PMMM; cytostatic;
KW immunomodulator; expression; immune disorder; cancer; human.
XX Homo sapiens.
OS
XX WO2004084806-A2.
XX 07-OCT-2004.
XX 16-MAR-2004; 2004WO-US0080006.
XX 21-MAR-2003; 2003US-0456864P.
PR 03-APR-2003; 2003US-0460512P.
PR 19-MAY-2003; 2003US-0472027P.
PR 30-MAY-2003; 2003US-0475072P.
PR 30-MAY-2003; 2003US-0475190P.
PR 24-JUN-2003; 2003US-0482575P.
PR 13-AUG-2003; 2003US-0495151P.
XX (INCY-) INCYTE CORP.
XX Baughn MR, Marquis JP, Kable AE, Chawla NK, Emerling BM, Lee SY;
PI Hafalia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura LL;
PI Yang J, Jin P, Becha SD, Wilson AD;
XX WPI: 2004-710256/69.
DR N-PSDB; ADT04072.
XX
XX New human protein modification and maintenance molecules (PMMM)
PT polypeptide, useful in preparing a composition for treating a disease
PT associated with decreased expression or overexpression of PMMM e.g.,
XX cancer.
XX Claim 1; SEQ ID NO 24; 222pp; English.
XX This invention relates to novel human protein modification and
CC maintenance molecule (PMMM) proteins and the DNA sequences which encode
CC them. The invention may be useful for the production of compounds with a
CC cytostatic or immunomodulator activity. The proteins are useful in
CC preparing a composition for diagnosing or treating a disease or condition
CC associated with decreased expression or overexpression of functional
CC PMMM, for example immune disorders or cancer. The present sequence is
CC that of a human protein modification and maintenance molecule (PMMM)
CC protein of the invention.
XX
SQ Sequence 738 AA;
Alignment Scores:
Pred. No.: 0 Length: 738
Score: 3817.00 Matches: 732
Percent Similarity: 83.1% Conservatives: 1
Best Local Similarity: 83.0% Mismatches: 5
Query Match: 68.8% Indels: 144
DB: 8 Gaps: 3
US-10-825-632-2 (1-3120) x ADT04044 (1-738)
QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATCACAGGATCGCCTTAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACAGAAAATATCATGGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACTCATTC 453

Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACCTTGCATCTCTGGTGAGAACAGAGAAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyLysArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCACAACTATCAATAGAGACAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGCACTATGGAATGATTCTCGAAGAAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACGATTTGGAACAGTCGGAATGCTCTTACGATTATCACCAAGGAAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGAGTAGTGGAAATTTATCAGCTAAAGATGGAGGCCACAAAGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCACAACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGAGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATGCCCCGCTGATCCAGACTGGATTCGCTTTTATATACATGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCACATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnGluLeu 240
QY 934 GCAACATCGAAAGAAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCCTGGCTATTGTTGGTGTCTCAAAAGCTGAAACAACTCCAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrSer ----- 276
QY 1054 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTATCATGTT 1113
Db 276 ----- 276
QY 1114 ACATCCCCTATTGTGGAACAACAGGAGGCGAGATTTCATTCGTTATCTCTAAACACAGTACA 1173
Db 277 -----Thr 277
QY 1174 GCMAATCTAAAGTCACCTTTAAGATGTCAGAAATATGATGATCTGCTGAAGAGGATC 1233
Db 278 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 297
QY 1234 ATGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCCTATTGAGAGGTTGAA 1293
Db 298 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 317
QY 1294 TATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTGGTCCATCTCTACTAGAT 1353
Db 318 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 337
QY 1354 CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTCACGAGTAA 1413
Db 338 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 357
QY 1414 GATGATGTTTATGGAAGGCAGAGACTCAATTGAGTCAGTGGCTGATTCCTGTGAGCCCACTA 1473
Db 358 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 377
QY 1474 ATTATCTTGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Db 378 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 397

QY 1534 CCCCAAAGTCACGAAGAGAAATTGAGTTTATTTTGCCTCTGAATCCAAAACAGGTTTC 1593
Db 398 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 417
QY 1594 CGTCATTATATACAAATATACATCTATTATTTAAAGGAAAGCAATATATAACCATCCAGTGGT 1653
Db 418 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGlyGly 437
QY 1654 GGCTGCTCCTCCCAAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTTACCACT 1713
Db 438 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 457
QY 1714 GGTGAATGGGAAGTCTTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773
Db 458 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 477
QY 1774 CTGCTATATTTTGAAGGACCAAGACTCCCTCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 478 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 497
QY 1834 TAGCTAAATCTCGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCTGC 1893
Db 498 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 517
QY 1894 ATCAGTCAGACTGTGACTCTCTTATTAAGTAAAGTATAGTACCAGAGAATCCACACTGT 1953
Db 518 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 537
QY 1954 GTGTCCCTTTACAGCTATCAGTCTCGAAGTACGCCCACTTCGCAAAAACAAAGGAATTT 2013
Db 538 ValSerLeuTyrLysLeuSerSerProGluAspAspProLeuAla ---AsnLysGluPhe 556
QY 2014 TGGGCCACATTTTGGATTCAGCAGGCTCTCTCTGACTACTCTCTCCAGAAATTTTC 2073
Db 557 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 576
QY 2074 TCTTTTGAAGTACTACTGGATTTACATGTATGGATGCTCTCAAGCCCTCATGACTA 2133
Db 577 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 596
QY 2134 CAGCTCGGAAGAAATATCTACTGCTGTTCATATATATGTTGCTCTCAGGTGACGTTG 2193
Db 597 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln ----- 613
QY 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGT 2253
Db 613 ----- 613
QY 2254 TATGTGTTGTAGTAGATAGACAAACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGGC 2313
Db 613 ----- 613
QY 2314 GCCTTTAAATATATAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATAT 2373
Db 613 ----- 613
QY 2374 CTAGCTTCTCGATATGATTTTCATTGACTGTAGATCGTGTGGGCATCCACGGCTGCTCTAT 2433
Db 613 ----- 613
QY 2434 GGAGGATACCTCTCCCTGATGTCATTAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATT 2493
Db 614 -----ValaIle 616
QY 2494 GCTGGGGCCCCAGTCTACTCTGTGGATCTTCTATGATCAGGATACCGAAGCTTATATG 2553
Db 617 AlaGlyProProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 636
QY 2554 GGTCAACCCTGACCAAGAAATGAAACAGGCTATTTACTTAGGATCTGTGGCCATCCCAAGCAGAA 2613
Db 637 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 656

QY 2614 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTC 2673
DB 657 LysPheProSerGluProAsnArgLeuLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 676
QY 2674 CATTTTGCACATACCAGTATATTACTAGCTTTTATAGTCAGGCTGGAAAGCCATATGAT 2733
DB 677 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 696
QY 2734 TTACAGATCTATCTCAGGAGAGACACAGCATTAAGATTCCCTGAATCGGAGAACATTAT 2793
DB 697 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 716
QY 2794 GAACCTGATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA 2853
DB 717 GluLeuHisLeuLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 736
QY 2854 GTGATA 2859
DB 737 Vallie 738
RESULT 14
ABB97362
ID ABB97362 standard; protein; 724 AA.
XX ABB97362;
AC ABB97362;
DT 27-JUN-2002 (first entry)
XX Novel human protein SEQ ID NO: 630.
DE Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
KW antinfertility; cerebroprotective; cytosatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag.
XX Homo sapiens.
OS WO200222660-A2.
XX 21-MAR-2002.
XX 10-SEP-2001; 2001WO-US026015.
XX 11-SEP-2000; 2000US-00659671.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-292408/33.
DR N-PSDB; ABB32548.
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
XX Example 2; SEQ ID NO 630; 509pp; English.
XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibit e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention
SQ Sequence 724 AA;

Alignment Scores:
Pred. No.: 0

Length: 724

Score: 3771.00 Matches: 724
Percent Similarity: 82.1% Conservative: 0
Best Local Similarity: 82.1% Mismatches: 0
Query Match: 67.9% Indels: 158
DB: 5 Gaps: 2
US-10-825-632-2 (1-3120) x ABB97362 (1-724)
QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTGCGGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATGAATACACAGATCGCGCTTAATATGGAGCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTGCGGATACACAGAAATATCATGCTACATCATG 393
DB 41 SerTyrSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGGCCATGCTCGTGAGAACAGAGAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTCGAAGCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAGCACTGCGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAA 633
DB 121 AspLeuPheGln----- 124
QY 634 GAAAGAAAACGCATTGGAAACAGTCGGAATGCTTCTTACGATTATACCAAGGAAGTGA 693
DB 124 ----- 124
QY 694 ACATTTCTGTTTCAAGCCGTTAGTGAATTTATCAGGTAAAGATGAGGGCCACAAGGA 753
DB 124 ----- 124
QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACACTAGTTGTGCCAACATACGATG 813
DB 125 -----GlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 142
QY 814 GATCCAAAATATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
DB 143 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 162
QY 874 TGATATCTAACATCGTAAACAGAGAGAAAGAGACACTCATTATGTGCACAAAGAGCTA 933
DB 163 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValIleAsnGluLeu 182
QY 934 GCCAACATGGAAGAAGATCCAGATCAGCTGAGGTGCTTACCTTTGTTTCTCCAAAGAGAA 993
DB 183 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 202
QY 994 TTTGATAGATATTTCTCGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGTGT 1053
DB 203 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 222
QY 1054 AAAATCTTAGAATCTTATATGAGAAATGATGATCTGAGGTGGAATTTATTCATGTT 1113
DB 223 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 242
QY 1114 ACATCCCTGATGTTGGAAACAGAGGGCAGATTCATTCCGTTTATCCCTTAAACACAGTACA 1173
DB 243 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 262
QY 1174 GCAAATCCTAAAGTCATCTTTTAAGATGTCAGAAATATGATGATGCTGCTGAAGGAAGATC 1233

Db	263	AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle	282	Db	599	-----	599
QY	1234	ATAGATGTCATAGATAGGAACTAAATCAACCTTTTGAGATCTCTATTTGAGGAGTTCGAA	1293	QY	2314	GCCTTTAAATATAAAATGGGTCAAAATAGAAATTCACGATCAGGTGGAGGACTCCAATAT	2373
Db	293	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	302	Db	599	-----	599
QY	1294	TATATTGCCAGAGCTGAGTGGACTCTCGAGGGAATAATGCTGGTCCATCTCTACTAGAT	1353	QY	2374	CTAGCTTCTCGATATGATTTTCATTGCACTTAGATCGTGGGCATCCACGGCTGGTCTTAT	2433
Db	303	TyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	322	Db	599	-----	599
QY	1354	CGCTCCAGACTCGCCCTACAGATAGTGTTCATCTCACTGAATATTATCCAGTAGAA	1413	QY	2434	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCCAGGGTTGCTATT	2493
Db	323	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	342	Db	600	-----	602
QY	1414	GATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTGCCTGATTCTGTGAGCGCACTA	1473	QY	2494	GCTGGGGCCCCAGTCACTCTCTGGATCTTCTATCATACAGATACACGGAAACGTTATATG	2553
Db	343	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	362	Db	603	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	622
QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533	QY	2554	GGTCAACCTCGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCCAAGCAGAA	2613
Db	363	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	382	Db	623	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	642
QY	1534	CCCCAAGCTCACGAGAGAAATTTAGTATTTTTCCTCTGATGCAAAACAGGTTTC	1593	QY	2614	AAAGTCCCTCTGGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2673
Db	393	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	402	Db	643	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	662
QY	1594	CGTCATTTATACAAATTTACATCTATTATTAAGGAAGCAATATATAACGATCCAGTGGT	1653	QY	2674	CATTTTGACATACACAGATATATTACTGAGTTTTTTTACTGAGGGCTGGAAGCCATATGAT	2733
Db	403	ArgHisLeuLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGly	422	Db	663	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	682
QY	1654	GGGCTGCTCTCCAGTATTTCAAGTGCCTCTCAAGAGAGAGATAGCAATTCACAGT	1713	QY	2734	TTACAGATCTATCCCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTAT	2793
Db	423	GlyLeuProAlaProSerAspPheLysCysProIleGlyGluIleAlaIleThrSer	442	Db	683	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	702
QY	1714	GGTGAATGGGAAGTCTTTGGCGGCATGATCTAATATCAAGTTGATGAAGTCAGAAGG	1773	QY	2794	GAACTGCATCTTTGCATCTCTCAGAACCTTCAGAACCTTCGATCAGTATTCGCTCTAAA	2853
Db	443	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	462	Db	703	GluLeuHisLeuLeuHisTyrLeuGlnGlySerArgIleAlaAlaLeuLys	722
QY	1774	CTGGTATATTTTGAAGGCACAAAGACTCCCTTTTAGAGCATCACCCTGACCTAGT	1833	QY	2854	GTGATA	2859
Db	463	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	482	Db	723	ValIle	724
QY	1834	TACGTAATCTCGAGAGGTGACAAAGGCTGACTGACCGGTGCTACTCACATTTCTTGTCG	1893	RESULT 15			
Db	483	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	502	ID	ABG61594	standard; protein; 690 AA.	
QY	1894	ATCAGTCAAGACTGTGACTCTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGT	1953	XX	ABG61594;		
Db	503	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	522	XX	12-AUG-2002 (first entry)		
QY	1954	GTGTCCTTTTACAGCTATCAAGTCTCGAAGATGACCCCACTTGCAAAACAGGAATTT	2013	XX	Human DRP-1 splice variant #1.		
Db	523	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	542	XX	Human; serine protease; dipeptidyl peptidase IV-related protein; DPP-IV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain; diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke; heart failure; hypertension; urinary retention; osteoporosis; cancer; ulcer; allergy; cancer; psychotic disorder; neurological disorder; dyskinnesia; reproductive disorder; inflammatory disorder; metabolic disorder.		
QY	2014	TGGGCCACATTTTGGATTCAGAGGTCTCTTCTCGACTATCTCTCCAGAAATTTTC	2073	XX	Homo sapiens.		
Db	543	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	562	XX	WO200231134-A2.		
QY	2074	TCTTTTGAAGTACTTACTGGATTTPACATTTGATGGGATGCTCTACAGCCTCATGACTA	2133	XX	18-APR-2002.		
Db	563	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	582	XX	12-OCT-2001; 2001WO-US031874.		
QY	2134	CAGCTTGAAAGAAATATCTCTACTGTCTCTTTCATATATGGTGGTCCCTCAGGTGAGTTG	2193	XX	12-OCT-2000; 2000US-0240117P.		
Db	583	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	599	XX	(FERR) FERRING BV.		
QY	2194	GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATATCCCTAGCTCTCTAGGT	2253	XX	Qi S, Akinsanya KO, Riviere PJ, Junien J;		
Db	599	-----	599	XX			
QY	2254	TATGTGTTGTAGTAGACAAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGC	2313	PI			

XX WPI; 2002-444178/47.
DR N-PSDB; ABR83325.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 59-61; 113pp; English.
PS
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPP-IV)-related proteins
CC (DPP-IV). The dipeptidyl peptidase IV-related proteins (DPP-IV) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypertension, hyperextension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPP-IV proteins
XX
SQ Sequence 690 AA;

Alignment Scores:
Pred. No.: 0 Length: 690
Score: 3607.50 Matches: 690
Percent Similarity: 87.1% Conservative: 0
Best Local Similarity: 87.1% Mismatches: 0
Query Match: 65.0% Indels: 102
DB: 5 Gaps: 1

US-10-825-632-2 (1-3120) x ABG61594 (1-690)

QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGAACTGGCGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGluLeuGlyValGluLeuPheGluThrAlaAspCys 20
QY 274 GAGGAGAAATTTGAATCATCAGGATCGGCTAAATTTGGAGGCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCTGAGTCAGCTTAAAGCTGCTTCCGATACCAGAAATATCATGCTCATCATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTTAGCCACCATGATTTTCATGTTGTTGAGAGGAATGATCCAGATGACCTCATTTCA 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAAATCTATTACCTTGCATGCTGCTGAGAACACAGAGAAATACACTGTTTATCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATCCCAAACTATCAATAGACAGCAGCTTATGCTCTCTTGGAGCCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCACACTGGACTATGAAATGATTTCTCGAGAGGAAGAACTATTAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCAATTCGAAACAGTCGGAATTTGCTTACGATTATCATCCAGGAGGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATCATACGTAAAGATGAGGGCCACAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCAATACGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATTGCCCCCTGATCCAGACTGGATTGCTTTTATACATAGCAACATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCTGTAAACAGAGAAAGAGAGACTCATTATGTGCACAAATGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAAATCGGAAGAGATCCAGATCAGCTGAGTCCGCTACTCTTTGTTCTCCAAAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGTAGATATTCTCGCTATTGCTGCTGCTCCAAAGCTGAAACAACTCCAGTGGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCTGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGAAACAGAGGGCGAGATTCTTCCGTTATCTTAAACAGAGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACTTTTAAGATGTCCAGAAATAATGATGTGCTGAAGAGGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCTATAGATAAGGAACATAATCAACTTTTGTAGATCTTATTTTGAAGAGTGA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATGCCAGACTGGATGGACTCTGAGGAGAAATAATGCTGTGCTCCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATCTTATGAAAGGAGAGACTCATCTAGTGCCTGCTGATCTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTCGATTAATATCCATGACATCTTTCATGTTTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCCAGAGGAGAAATGAGTTTATTTTGTGCTGAAATGCAAAACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATATAACGATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGTGTCTCTCCAGTGAATTCAGTGTCTCTATCAAGAGAGATAGCATTAACCAT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAGTCTTCTGCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAAGAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGTGTATTTTGAAGCCACAAAGACTCCCTCTTAGAGCATCACCTGTACGTAGTACGT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACGTAATCTCGAGAGGTGACAGGCTGACTGACCGGTGCTACTCAATCTCTGTGCTGC 1893

Qy	430	AATGATCCAGATGACCTCATTACAGACAGAACTATTACCTTGCCATCTCTGGTGAGAAC	489	Qy	1288	-----GTTGAATATATTGCCAGAGCTGGATGACTCCTCGAGGAAATATGCTTGGTCC	1341
Db	73	MetArgAlaGlu-----LysAspArgLeuAsnMetTyrAlaIleSerSerValPro	89	Db	388	ProPheTyrGluTyrIleThrArgAlaGlyPhePheSerAspGlyThrThrValTyrVal	407
Qy	490	AGAGAAATACA-----CTGTTTATTCTGMAATATCCC-----AAAACCTATCAATAGAGCA	540	Qy	1342	ATCCTACTAGATCGCTCCAGACTCGCTCATAGATGTTGATC-----	1386
Db	90	GlyThrAsnThrGlnSerIlePheSerValThrIleProLeuGluLeuValGluLysAla	109	Db	408	GlnValMetSerArgAspGlnAlaGlnCysSerLeuLeuLeuIleProTyrThrAspPhe	427
Qy	541	GCAGCTTAAGTCTCTCTGGAAGCCTCTTTTGGATCTTTTCAGGCAACACTCGACTAT	600	Qy	1387	-----TCACCTGAA	1395
Db	110	GlnValAlaAspArgLys-----PheGluLeuLysLeuLysSer	122	Db	428	LeuLeuProGluGluLeuGlyGlySerIleLysLeuAspAsnLeuLysLeuSerThrAsp	447
Qy	601	GGAATGTAT-----TCTCGAGAA	618	Qy	1396	TTATTATCCAGTAGAAGATGATCTTATGGAAGCAGAGACTCATTTAGTCAGTGCCT	1455
Db	123	GlyTyrAsnValAspSerTyrIleArgMetSerCysArgLysThrProProSerAlaGlu	142	Db	448	LeuAsnMetGlyValTyrAspAsp---LysSerHisGluGluThrMetGluLysProPro	466
Qy	619	GAAGACTATTAAAGAGAGAAACCAATGGACAGTC---GGAATGCTCTCTACCAT	675	Qy	1456	GATTCTGTGAGC-----CCACTAATATCTATGAAGAAAACAACAGACATCTCGATAAAT	1509
Db	143	PheThrLeuGlnCysGluArgGlnArgSerGlnValValThrGlyIleSerAspTyrGlu	162	Db	467	ArgGlyLysLeuArgGlyThrValGlnIleHisLysAlaArgAsnAspTyrTyrIleAsn	486
Qy	676	TATCACCAAGGA-----AGTGGAACTATCTCTGTTTCAAGCCGT-----	714	Qy	1510	ATCCATGATCATCTTTCATGTTTTTCCCAAGTCACGAAGAGAA-----ATTGAGTTT	1563
Db	163	IleArgAsnGlyLysMetIleLeuMetAlaGlyAspGlnLeuPheArgTyrAsnProLeu	182	Db	487	ThrHisAsnAlaIleTyrProLeuLysIleThrAspGluGluHisProMetTyrGluPhe	506
Qy	714	-----	714	Qy	1564	ATTTTGCTCTGATGCAAAACAGGTTTCGTCATTATATACAAAATATCATCTATTATA	1623
Db	183	AsnGluAlaLeuAlaAlaIleProIleAlaValProAspAspGlnSerSerThrGluPro	202	Db	507	IleTyrCysLeuGluLysProAsnGly---SerCysLeuAlaLeuIleSerAlaGluLeu	525
Qy	715	-----AGTGGAAATTTATCAGCTAAAGATCGAGGCCA	747	Qy	1624	AAGGAAGCAATATAACGATCCAGTGGTGGCTGCTCTCCAAGTATTCAAGTGT	1683
Db	203	MetAspIleSerGluGlySerIleThrSerGlyThrLysGlyCysSerAsnGluAlaPro	222	Db	526	AspGlnAsnGlyTyr-----Cys	531
Qy	748	CAAGGA-----TTTACGCAACAACTTTTAAAGGCCCAATCTAGTGGAACT	792	Qy	1684	CCTATCAAAAGAGAGATAGCAATACACGCTGCTGAATGGGAAGTTCTTGCCCGCATGGA	1743
Db	223	GlnSerSerThrValProValThrArgIleProIleLysLysProThrThrSerThr	242	Db	532	ArgHisThrGluGluLysLeuLeuMetAlaGluAsnPheSerIle-----AsnLysSer	549
Qy	793	AGTTCTGCC-----AACACGCGATGGATCCAAATATATGCCCC	831	Qy	1744	TCTAATATCAAGTTGATGAAGCTCAGAAAGCTGCTATATTTTGGAGCCACCAAGACTCC	1803
Db	243	GluLysProAlaThrAlaProProThrAsnAsnPheValSerSerAlaLysValCysPro	262	Db	550	MetGlyIleValValAspGluValArgGluLeuValTyrTyrValAlaAsnGluSerHis	569
Qy	832	GCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTA	891	Qy	1804	CCTTTAGACATCACCCTGTAGCTAGTACGTAACTCTGAGAGGTGACAGGCTG	1863
Db	263	AlaAspSerSerLeuLeuAlaTyrValLeuAsnLysGlnValTyrIle-----	278	Db	570	ProThrGluTyrAsnIle---CysValSerHisTyrArgThrGlyGlnHisAlaGlnLeu	588
Qy	892	ACCAGAGAAGAGGAGACTCACATTATGTCACAAATGAGCTAGCCACATGGAAGAAGAT	951	Qy	1864	ACTGACCGTGGCTACTCACAATCTTCTGCTGCATCAGTCAGCAGCTGTCTTTTAAAGT	1923
Db	279	-----GluLysAsnGlyLysIleIleHisArgThrSerSerAsnSerLysHisIle	295	Db	589	ThrGluSerGlyIle-----Cys	597
Qy	952	GCAGATCAGCTGGAGTCCTACCTTGTCTCCAGAGAATTTGATAGATATCTGCG	1011	Qy	1924	AAGTATAGTAACACAGAAG-----AATCCACACTGTGTCTCTCTTTACAAGCTA	1971
Db	296	ThrAsn-----GlyValProSerTyrIleValGlnGluLeuGluArgPheGluGly	313	Db	598	GluArgAlaAsnGlyLysLeuAlaLeuAspLeuAspHisGlyPheAlaCysTyrMetThr	617
Qy	1012	TATGTGGTGTCCAAAAGCTGAAACAACTCCAGCTGGTGGTAAATTTCTTAGAATCTA	1071	Qy	1972	TCA-----AGTCCTGAAGATGACCAACTTGCAAAACAAAGGAATTT---TGG-----	2016
Db	314	IleTyrTrp---SerGluSerLysThr-----ArgLeuLeu	324	Db	618	SerValGlySerProAlaGlu-----CysArgPheTyrSerPheArgTyrLysGlu	634
Qy	1072	TATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT-----	1113	Qy	2017	-----GCCACATTTTGGATTTCAGCAGGTCTCT---	2043
Db	325	TyrGluHisValAsnGluLysValAlaGluSerGlnPheGlyValAsnGlyAspPro	344	Db	635	AsnGluValLeuProSerThrValTyrAlaAlaAsnIleThrValSerGlyHisProGly	654
Qy	1114	---ACATCCCTCTATGTTGGAAACAGGAGGCGAGATTCAATCCGCTTATCTCTTAAACAGGT	1170	Qy	2044	CTTCTCTGAC-----TATACTCTCCAGAAATTTCTCTTTTGAAGT---ACTACTGGA	2094
Db	345	ProValAlaProMet-----LysTyrProArgAlaGly	355	Db	655	GlnProAspLysHisPheAspSerProGluMetIleGluPheGlnSerLysLysThrGly	674
Qy	1171	ACAGCAAAATCTTAAAGTCACTTTTAAGATGTGAGAAATAATGATGCTGAGGAAGG	1230	Qy	2095	TTTACATTTGATGGGATGCTCTACAGCTCATGATCTACAGCTGGAAGAAATATCCT	2154
Db	356	ThrLysAsnAlaTyrSerThrLeuArgMet-----ValIleLeuGluAsnGlyLys	372	Db	675	LeuMetHisTyrAlaMetIleLeuArgProSerAsnPheAspProTyrLysLysTyrPro	694
Qy	1231	ATCATAGATGTCATAGATAAGCACTAATTCACCTTTTGGATTTCTATTATTAAGGA---	1287	Qy	2155	ACTGTGCTGTTCATATATGCTGCTCTCAGGTGCGATGCTGGAATATCGGTTTAAAGGA	2214
Db	373	AlaTyrAspValProLeuLysAsp-----GluValIleTyrLysHisCys	387	Db	695	ValPheHisTyrValTyrGlyGlyProGlyIleGlnIleValHisAsnAspSerTrp	714
Qy				Qy	2215	GTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTATGCTGTTAGTATAGAC	2274

Ds	715	ileInTyrllearg-----PheCysArgLeuGlyTyrValValPheIleAsp	731						Qy	436	CCAGATGGACCTCAITTCAGACAGAACTTATACCTTGCATCTGTGTGAGAACAGAGAA	495
Qy	2275	AACAGGGGATCTGTCCACGAGGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGT	2334						Ds	47	ProaspGly-----LysArgValThrTyrLeuLysGlyLysProGlu-----	60
Ds	732	AsnArgGlySerAlaHisArgGlyIleGluPheGluArgHisIleHisLysLysMetGly	751						Qy	496	AATACACTGTTTTTATCTGAAATTCCTCAAACTATCAATAGAGCAGCAGCTTCTTAATGCTC	555
Qy	2335	CAATAGAAATTCAGCATCAGGTGGAAGACTCCCAATATCTAGCTTCTCGATAT- --GAT	2391						Ds	61	-----AlaAlaAsnIleGlnAsp	66
Ds	752	ThrValGluValGluAspGlnValGluGlyLeuGlnMetLeuAlaGluArgThrGlyGly	771						Qy	556	TCITGG-----AAGCCTCTTTTGGATCTTTTTCAGGCCAACACTG	594
Qy	2392	TTCAATGACTTAGTGTGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTG	2451						Ds	67	LeuTrpAlaAlaAspValLysGlyGluProTyrArgLeuIleAspSerAlaAlaLeu	86
Ds	772	PheMetAspMetSerArgValValValHisGlyTrpSerTyrGlyTyrMetAlaLeu	791						Qy	595	GACTATGGA-----ATGTATTCTCGAAGAGAACTATTAAAGAGAAAGAAACGCATT	648
Qy	2452	ATGGCATTATGACAGAGTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCGCCACT	2511						Ds	87	SerSerGlyAspLysGluLeuSerGluAlaGluLysAlaArgGluArgAlaArgVal	106
Ds	792	GlnMetIleAlaLysHisProAsnIleTyrArgAlaAlaIleAlaGlyAlaValSer	811						Qy	649	GGAAACAGTCGGAATTCCTTACGATTATCACCAAGGAAGTGGAAACATTTCTGTTTCAA	708
Qy	2512	CTGTGGATCTTCATGATACAGATACACGGAACGTTATATGGGTCCACCTGACCCAGAAAT	2571						Ds	107	SerAlaArgGlyIleValGluTyrSerTrpAspArgGlnGlyArgPheIleLeuValPro	126
Ds	812	AspTrpArgLeuTyrAspThrAlaTyrThrGluArgTyrMetGlyTyrPro- --LeuGlu	830						Qy	709	GCCGTAGTGAATTTAT-----CACGTAAAGATGAGGGCCACCAAGGATTTACGCCAA	762
Qy	2572	GAACAGGCTTATCTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCA	2631						Ds	127	LeuAspGlyAspLeuTyrLeuAspAlaValAlaAspGlyLysIleThrArgLeuThrGlu	146
Ds	831	GluHisValTyrGlyAlaSerSerIleThrGlyLeuValGluLysLeuProAspGluPro	850						Qy	763	CAACCTTTAAGGCCCAATCTAGTGGAACTAGTTCTCCCAACATACGATCGATGCCAAA	822
Qy	2632	AATCGTTTACTCTTACATGCTTTCTCGATGAGATGCAATTTGTCACATACACAGT	2691						Ds	147	-----ThrProGlyAspGluValAspAlaLys	155
Ds	851	AsnArgLeuMetLeuValHisGlyLeuMetAspGluAsnValHisPheAlaHisLeuThr	870						Qy	823	TTATGCCCGCTGATCCAGACTGGATGCTTTTATACATACCAACGATTTTGGATCT	882
Qy	2692	ATATTACTAGTTTTTACTGAGGCTGGAAGCCATATGATTACAGATCTATCCTCAG	2751						Ds	156	ValSerProLysGly- --GlyTyrValSerTyrValArgAspGlnAsnLeuTyrIleLys	174
Ds	871	HisLeuValAspGluCysIleLysGlyLysTyrTrpHisGluLeuValIlePheProAsn	890						Qy	883	ACATCGTAAACACAGAGAAAGGAGACTCACTATGTGCAATGACAGCTAGCCCAACATG	942
Qy	2752	GAGACACAGCAGCAAGAGTCTCGAATCGGAGAACATTTATGAACTGTCATCTTTGCGAC	2811						Ds	175	ProValAlaGlyGlyAlaGluThrAlaLeuThr-----ThrAsp	187
Ds	891	GluArgHisGlyValArgAsnAsnAspAlaSerIleTyrLeuAspAlaArgMetMetTyr	910						Qy	943	GAAGAGATGCCAGATCACCTGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGA	1002
Qy	2812	TACCTTCAAGAA	2823						Ds	188	GlyLysAspAlaLeuSerPheGlyValAlaGluPheIleValGlnGluLeuAspArg	207
Ds	911	PheAlaGlnGln	914						Qy	1003	TATTCGGCTTATGTGGTGTCTCAAAAGCTGAAACAACTCCACAGTGGTGGTAAAAATCTT	1062
RESULT 2									Ds	208	PheThrGlyTyrTrpTrpSerProAspGluSer-----	218
A87516									Qy	1063	AGAATCTTATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCT	1122
C:Species: Caulobacter crescentus									Ds	219	ArgIleValTyrThrArgValAspGluSerGlyValAspIleVal-----	233
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004									Qy	1123	ATGTTGGAACAAAGGGGCGAGAT-----	1155
C:Accession: A87516									Ds	234	-----ProArgAlaAspIleGlyProGlyGlyAlaThrValValAsnGlnArg	249
R:Nierman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001									Qy	1156	TATCCTAAACAGGTACACAAATCTTAAAGTCACCTTTTAAGATGTGCAGAAATAATGATT	1215
A:Title: Complete Genome Sequence of Caulobacter crescentus.									Ds	250	TyrProArgAlaGlyArgProAsnAlaValValAspLeuPheValArgAspLeu-----	267
A:Reference number: A87249; MUID:21173698; PMID:11259647									Qy	1216	GATGCTGAAGGAAGGATC-----ATAGATGTC-----ATAGATAAGGAACATAATCAACCT	1266
A:Accession: A87516									Ds	268	--AlaSerGlyLysValThrAlaLeuAspLeuGlyAlaAsnLysAspIle-----	283
A:Molecule type: DNA									Qy	1267	TTTGAGATCTATTGGAAGAGTTGAATATATTTGCGAGAGCTGGATGGACTCTCTGAGGA	1326
A:Residues: 1-738 <STO>									Ds	284	-----TyrValAlaArgValAlaTrpSerAlaAspGly	294
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:gl3423647; E C:Genetics: CC2154									Qy	1327	AAATATGCTTGGTCCATCTCTACTAGATCGCTCCCGACTCGCTTACAGATAGTGTTCATC	1386
Alignment Scores:									Ds	295	LysThrValTyrValGlnArgLeuSerArgAspGlnLysThrLeuAspLeuLeuAlaPhe	314
Pred. No.: 5,93e-53									Qy	1387	TCACCTGAATTTATCCAGTAGAAGATGATGTTATGGAAAGGCAGAGACTCATTTGAG	1446
Score: 842.00									Ds	314	-----	314
Percent Similarity: 43.9%									Qy	1447	TCAGTGCCTGATTTCTGTGACGCCACTA-----ATTATCTATGAAGAAACAAACAGACATC	1500
Best Local Similarity: 29.2%												
Query Match: 15.2%												
Indels: 170												
Gaps: 2												

US-10-825-632-2 (1-3120) x A87516 (1-738)

Db 129 GlyGluLeuTyrLeuTyrAspLeuLysGlnGluGlyLysAlaAlaValArgGlnLeu--- 147
QY 769 TTAAGGCCAATCTAGTGGAACTAGTTGTGCCAACATACGCGATGGATCCAAATATTATGC 828
Db 148 -----ThrHisGlyGluGlyPheAlaThrAspAlaLysLeuSer 160
QY 829 CCCCTGATCCAGCTGGATTGCTTTTATACATAGCAAGCATATTTGGATATCTAACATC 888
Db 161 ProlysGly---GlyPheValSerPheIleArgGlyArgAsnLeuTrpValIleAspLeu 179
QY 889 GTAACACAGAGAAGAGACACTCTATGTGCGCAATGAGCTAGCCCAACATCGAAGAA 948
Db 180 AlaserGlyArgGlnMetGlnLeuThr-----Ala 189
QY 949 GATCCAGATCAGCT-----GGAGTCGCTACCTTTGTTCTCCAAGAAGATTTGAT 999
Db 190 AspGlySerThrThrIleGlyAenGlyIleAlaGluPheValAlaAspGluuWetAsp 209
QY 1000 AGATATTCTGGCTATTGGTGTCTCCAAAGCTGAAACAACTCCCGAGTGGTGTAATAAT 1059
Db 210 ArgHisThrGlyTyrTrpAlaProAspAspSerAla----- 222
QY 1060 CTTAGATTCTATATGAGAAATGATGATCTGAGTGC-----GAA 1101
Db 223 -----IleAlaTyrAlaArgIleAspGluSerProValProValGlnLysArgTyrGlu 240
QY 1102 ATTATTCTGATCATCCCTATGTTGGAAACAAGGAGGCGCAGATTTCATTCCTGTTATCCT 1161
Db 241 ValTyrAlaAspArgThrAspValIleGluGln-----ArgTyrPro 254
QY 1162 AAAACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTGCAGAAATAATGATGATGC 1221
Db 255 AlaAlaGlyAspAlaAsnValGlnValLysLeuGlyValIleSerProAlaGluGlnAla 274
QY 1222 GAAGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATCTATTT 1281
Db 275 GlnThrGlnTrpIleAspLeuGlyLysGluGlnAspIle----- 287
QY 1282 GAAGGAGTTGAATATATGCGCAGCTGGATGAGCTCCTGAGGGAAATATGCTTGCTCC 1341
Db 288 -----TyrLeuAlaArgValAsnTrpArgAspProGlnHisLeuSerPheGln 303
QY 1342 ATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACTCGAATATTT 1401
Db 304 ---ArgGlnSerArgAspGlnLysLysLeuAspLeuValGluValThrLeu----- 319
QY 1402 ATCCCAAGTAGATGATGTATGGAAGGCGAGAGACTCATTTGAGTCAGTCCGCTGATTCT 1461
Db 320 -----AlaSerAsnGlnArg----- 325
QY 1462 GTGAGCCCAATATTATCATGAAGAAACAACAGACATCTGGATAAATATCATGCATC 1521
Db 326 -----ValLeuAlaHisGluThrSerProThrTrpValProLeuHisAsnSer 341
QY 1522 TTTTCATGTTTTCCTCCCAAGTCCAGGAAGAGAAATGAGTTTATTTTGGCTCTGGAATGC 1581
Db 342 LeuArgPheLeu-----AspAspGlySerIleLeuTrpSerSerGlu--- 355
QY 1582 AAAACAGGTTTCCGTCTATATACAAAATTTACATCTATTTTAAAGGAAGCAATATATAA 1641
Db 356 ArgThrGlyPheGlnHisLeuTyrArgIle-----AspSerLysGlyLys 370
QY 1642 CGATCCAGTGGTGGCTGCTCTCCCAAGTATTCATGATGTCCTATCAAGAGAGATATA 1701
Db 371 AlaAla----- 372
QY 1702 GCAATTACCACTGTAATGGGAAGTTCTTGGCGCGCATGATCTAATATCAAGATTGAT 1761
Db 373 AlaLeuThrHisGlyAsnTrpSerVal-----AspGluLeuLeuAlaValAsp 388
QY 1762 GAAGTCAGAAGGCTGGTATATTTTGAAGGCGCAACCAAGACTCCCTCTTTAGACATCACCTG 1821
Db 389 GluLysAlaGlyLeuAlaTyrPheArgAlaGlyIleGluSerAlaArgGluSerGlnIle 408

RESULT 4

S66261

X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum

C:Species: Flavobacterium meningosepticum

QY 1822 TACCTAGTCAGTACGTAAATCTCGAGAGGTGACAAAGCTGACGTACGCGGTCTACTCA 1881
Db 409 TyrAlaValPro---LeuGlnGlyGlyGlnProGlnArgLeuSerLysAlaProGlyMet 427
QY 1882 CATCTTCTGCTGCATCAGTCAGCTAGCTGCTCTTTATAGTAAGTAACTATATAGTACAGAAG 1941
Db 428 HisSerAlaSerPheAlaArgAsnAlaSerValTyrValAspSerTrpSerAsnAsnSer 447
QY 1942 AATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAA 2001
Db 448 ThrProGlnIleGluLeuPheArgAlaAsnGly----- 459
QY 2002 ACAAAGGAATTTTGGGCGCACCATTTTG-----GATTCAGCAGCTCTCTCTCCTGACTAT 2055
Db 460 -----GluLysIleAlaThrLeuValGluAsnAspLeuAlaAspProLysHisProTyr 477
QY 2056 ACT-----CCTCCAGAAATTTCTCTTTTGAAGTACTACTGGA 2094
Db 478 AlaArgTyrArgGluAlaGlnArgProValGluPheGlyThrLeuThrAlaAlaAspGly 497
QY 2095 TTTACA---TTGTATGGATGCTCTACAAGCTCATGATCTACAGCTCGAAAGAAATAT 2151
Db 498 LysThrProLeuAsnTyrSerValIleLysProAlaGlyPheAspProAlaLysArgTyr 517
QY 2152 CCTACTGTGCTGTTCATATATGTTGCTCCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAA 2211
Db 518 ProValAlaValTyrValTyrGlyGlyProAlaSerGlnThrValThrAspSerTrpPro 537
QY 2212 GGAATCAAGTATTTCCGCTTGAATACC---CTAGCTCTCTAGGTATGTTGTTGTAGTG 2268
Db 538 GlyArgGlyAspHisLeuPheAsnGlnTyrLeuAlaGlnGlnGlyTyrValValPheSer 557
QY 2269 ATACACAACAGGGGATCCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAA 2328
Db 558 LeuAspAsnArgGlyThrProArgArgGlyArgAspPheGlyGlyAlaLeuTyrGlyLys 577
QY 2329 ATGGGTCAAATAGAAAATTGACGATCAGGTGGAAGGACTTCAATATCTAGCTTCTCGATAT 2388
Db 578 GlnGlyThrValGluValAlaAspGlnLeuArgGlyValAlaTrpLeu---LysGlnGln 596
QY 2389 GATTTCATGACTAGATCGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCC 2448
Db 597 ProTrpValAspProAlaArgIleGlyValGlnGlyTrpSerAsnGlyGlyTyrMetThr 616
QY 2449 CTGATGCATTAATGACAGGTCAGATATCTTCAGGTTGCTATTTGCTGGGCGCCAGTC 2508
Db 617 LeuMetLeuLeuAlaLysAlaSerAspSerTyrAlaCysGlyValAlaGlyAlaProVal 636
QY 2509 ACTCTGTGGATCTTCTATGATACAGGATACCGAAACGTTATATGGTCACTGACCAG 2568
Db 637 ThrAspTrpGlyLeuTyrAspSerHisTyrThrGluArgTyrMetAspLeuProAlaArg 656
QY 2569 AATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTCCCTCTGAA 2628
Db 657 AsnAspAlaGlyTyrArgGluAlaArgValLeuThrHisIleGluGlyLeuArgSerPro 676
QY 2629 CCAAATCGTTTACTGCTCTTACATGATGTTTCTGGATGAGAATGTCATTTTGCACATACC 2688
Db 677 -----LeuLeuLeuIleHisGlyMetAlaAspAspAsnValLeuPheThrAsnSer 693
QY 2689 AGTATATTACTGATGTTTTTAGTGGCGCTGGAAAGCCATATGATTTACAGATCTCTCCT 2748
Db 694 ThrSerLeuMetSerAlaLeuGlnLysArgGlyGlnProPheGluLeuMetThrTyrPro 713
QY 2749 CAGGAGACACACATCAAGATTCCTGAATCGGAGAACATTTATGAACCTG 2799
Db 714 GlyAlaLysHisGlyLeuSerGlyAlaAspAlaLeuHisArgTyrArgVal 730

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S66261
R:Kabashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A:Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from *Fluorobacterium*
A:Reference number: S66261; MUID:95314307; PMID:7793970
A:Accession: S66261
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-711 (KAB)
A:Cross-references: UNIPROT:Q47900; UNIPARC:UPT0000082115; EMBL:D42121; NID:G577283; PIR:1000000000
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase

Alignment Scores:
Pred. No.: 7,38e-37 Length: 711
Score: 620.00 Matches: 195
Percent Similarity: 42.6% Conservative: 134
Best Local Similarity: 25.2% Mismatches: 284
Query Match: 11.2% Indels: 160
DB: 2 Gaps: 26

US-10-825-632-2 (1-3120) x S66261 (1-711)
646 ATGGACAGTCGGAATTCCTTACGATTATCACCAGGA----- 687
54 IIEGLUPROTHRGlyIleAlaLysTyrSerTyrLysThrSerGlnLysGluLysAsnIle 73
688 ----AGTGGACATTT-----CTGTTTCAA 708
74 ValAspGlySerPheGlnGlyTyrThrPheSerAsnAspGluSerLysIleLeuLeuGln 93
709 GCGGTAGT-----GGATTTATCAGCTAAAGATGA 741
94 LysSerSerGlnSerIleTyrArgHisSerPheLeuGlyLysPheGluValLysAsp--- 112
742 GGGCCACAAGGATTTACGCAACACCTTTAAGGCC---AATCTAGTGGAACTAGTTCT 798
113 -----LeuLysSerArgThrValValSerLeuAsn 123
799 CCAACATACGGATGGATCCAAATATATGCCCCGCTGATCCAGCTGGATGCTTTATA 858
124 AlaAsnTrpIleGlnLysPheSerPro---AspGlySerLysValAlaPheIle 142
859 CATAGCACCATATTTGGATATCTAACATCGTAAACAGAGAAAGAGACTCATTTAT 918
143 AlaAspAsnAsnLeuPheTyrGlnAspLeuAsnThrGlyLysIleThrGlnIleThr 162
919 -----GTGCACATAGCTAGTACCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCT 972
163 AspGlyLysLysAsnGluIleIleAsn-----GlyLeuGly 174
973 ACCTTTCTTCCCAAGAAGATTTGATAGATATCTGGCTATTGGTGGTGTCCAAAAGCT 1032
175 AspTrpValTyrGluGluGluPheGlyHieAlaAspTyrTyrGlnTrp----- 190
1033 GAACAACATCCAGTGGTGTAAATCTTACGAATCTATATATGAAGAAATGATGATCT 1092
191 -----AsnLysAlaGlyAspAlaLeuValPheValArgPheAspGluArgLysValPro 208
1093 GAGTGGGAATTT---ATTGATGTATCC-----CCTATGTCGAACAGAGAGGCA 1143
209 GluIleAsnIleProIleTyrTyrGlnAsnLeuTyrProLysLeuMetThr----- 225
1144 GATTCATTCCTTATCTTAAACAGAGTACAGCAATCTTAAGTCTCTTTAAGATGCA 1203
226 -----TyrLysTyrProLysAlaGlyGluGluAsnSerAlaValThrAlaTyrLeuTyr 243
1204 GAAATATATGATGTGATGAGGAGGAGGATCATAGATGTCTATAGTAAGAACTAATCAA 1263
244 GlnLeu-----SerSerGlyLys----- 249
1264 CCTTTTGAGATTCATTATTAGAGGATTGAA-----TATATTGCCAGAGCTGGAGT 1317

250 SerAlaGlnLeuAsnPheGlySerSerGluLysTyrTyrIleProGlnLeuPheGlnThr 269
1318 CCTGAGGAAATATGCTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATA 1377
270 AsnAlaAsnAspGluLeuValValAlaThrAlaAsnArgHisGlnAsnLysValAspLeu 289
1378 GTGTGTGATCTCAGCTCGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGCA 1437
290 LeuLysValAsnThrLys----- 295
1438 CTCATTGAGTCAGTCCGCTGATCTCTGAGCCCACTAATTTATCTATGAGAACAACACAGAC 1497
296 -----ThrAlaAlaValSerLysLeuPheThrGluThrAspAsn 308
1498 ATCTGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCCAGAGAGGAATTT 1557
309 AlaTrpIleGluThrAspAsnLeuThrMetGluPhe-----LeuAspAspAsn 324
1558 GAGTTTATTTTGGCTCTGATGCAAAACAGGTTTCCGTCATTATPACAAAATTCATCT 1617
325 SerPheLeuTrpAlaSerGlu---ArgAspGlyHisArgHisLeuTyrTrpTyrAspAla 343
1618 ATTTTAAAGGAAGCAAAATATAACGATCCAGTGGTGGCTGCTCTCCAGTATTTC 1677
344 AlaGlyLysLeuLysLys----- 349
1678 AAGTGTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGG 1737
350 -----GlnValSerLysGlyAspTrpGluIleIleAsnTyr 361
1738 CATGATCTATATATCAAGTTGATGAGTCAGAGGCTGATATTTTGAAGGACCAAAA 1797
362 TyrGlyTyrAsnProLysThrLysGlu-----ValTyrIleGlnThrThrGlu 377
1798 GACTCCCTTTTAGACATCACCCTGTCAGTCTAGTCTAGT-----TACGTAAATCTGGAGAG 1851
378 LysGlySerLysAsnLys-----ValValSerLysLeuAsnIleAsnThrGlyLys 394
1852 GTGACAAAGCTGACTGACCGGTGCTACATCTTCTGTCATCATGCTGACACTGTGAC 1911
395 ThrGlnLeuLeuSerAsnAlaGluGlyAsnAsnSerAlaAlaPheSerLysThrPheAsn 414
1912 TTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCTTACAGCTA 1971
415 TyrPheIleAsnThrSerSerThrAlaLysValProThrLysTyrIleLeuLysAspAla 434
1972 TCAAGTCTCGAA-----GATGACCAACTTGCACAAACAAAGGAA 2010
435 AsnGlyLysAspValLysGluLeuGlnAsnAsnAspAspLeuLeuAsnLysLeuLysSer 454
2011 TTTTGGGCCCACTTTTGGATTTCAGAGGCTCTCTCTCTGACTATATCTCTCCAGAAAT 2070
455 -----AspAsnPheIleAlaLysGluPhe 462
2071 TTCTCTTTGAAAGTACTACTGATTTTACATTGTATGGATGCTCTACAGCCCTCATGAT 2130
463 IleThrIleProAsnAlaAlaGlyAspGlnMetAsnAlaTrpMetIleLysProLysAsn 482
2131 CTACAGCTGGAAGAAATATCTCTGCTGCTGTTCATATATGTTGGTCTCTCAGGTGAG 2190
483 PheAspProAlaLysLysTyrProValPheMetPheGlnTyrSerGlyProGlySerGln 502
2191 TTGTGGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTCAATACCTAGCTCTCTA 2250
503 GlnValAlaAsnSerTrpAspGlyGlyAsnGlyIleTrpPheAspMetLeuAlaGlnLys 522
2251 GGTATGTGGTGTGATGATAGACACAGGGGATCTCTGTCACCGAGGCTTAAATTCGA 2310
523 GlyTyrLeuValValCysValAspGlyArgGlyThrGlyPheArgGlyThrLysTyrLys 542
2311 GSGCGCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGTGGAGGACTCCAA 2370

Db 543 LysValThrTyrLysAsnLeuGlyLysTyrGluIleGluAspGlnIleThrAlaAlaLys 562
QY 2371 TATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATCCAGCGTGTGCC 2430
Db 563 TrpLeuGlyAsnGln---SerTyrValAspLysSerArgIleGlyIlePheGlyTrpSer 581
QY 2431 TATGAGGATCACTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGTTGGCT 2490
Db 582 TyrGlyGlyTyrMetAlaSerLeuAlaMetThrLysGlyAlaAspValPheLysMetGly 601
QY 2491 ATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGAACTGTTAT 2550
Db 602 IleAlaValAlaProValThrAsnTrpArgPheTyrAspSerIleTyrThrGluArgPhe 621
QY 2551 ATGGGTCACCTGCACGAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATCCAGCA 2610
Db 622 LeuGlnThrProGlnGluAsnLysAspGlyTyrAspLeuAsnSerProThrTyrAla 641
QY 2611 GAAAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGTTCTGCTGTTCTGATGAGAAT 2670
Db 642 LysLeuLeu-----LysGlyLysPheLeuLeuIleHisGlyThrAlaAspAsn 658
QY 2671 GTCCATTTTCACATACCAGTATATTACTGAGTTTCTAGTGGGCTGGAGCCATAT 2730
Db 659 ValHisPheGlnAsnSerMetGluPheSerGluAlaLeuIleGlnAsnLysLysGlnPhe 678
QY 2731 GATTTCACAGATCTATCTCTCAGGAGACACAGATAGAGTTCTCTGAATCGGAGAACAT 2790
Db 679 AspPheMetAlaTyrProAspLysAsnHisSerIleIleGlyGlyAsnThrArgProGln 698
QY 2791 TATGAACGTGATCTTTTGCACTACCTTCAAGAAACCTT 2829
Db 699 LeuTyrGluLysMetThrAsnTyrIleLeuGluAsnLeu 711
RESULT 5
CDHU26
dipeptidyl-peptidase IV (EC 3.4.14.5) - human
N/Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004
C/Accession: S24313; A42408; B61136; S59510; I56154; S59857; S15520
R/Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehara, Y.
Biochim. Biophys. Acta 1131, 333-336, 1992
A/Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a serine protease
A/Reference number: S24313; MUID:92329551; PMID:1352704
A/Accession: S24313
A/Molecule type: mRNA
A/Residues: 1-6,'I',8-766 <MIS>
A/Cross-references: UNIPROT:P27487; UNIPARC:UPI000016A880; EMBL:X60708; NID:g35335; PIDN:R;Datmoul, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A. J. Biol. Chem. 267, 4824-4833, 1992
A/Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer cell lines during cell differentiation.
IV mRNA levels during cell differentiation.
A/Reference number: A42408; MUID:92165847; PMID:1347043
A/Accession: B42408
A/Molecule type: mRNA
A/Residues: 1-5,'R',7-436,'S',438-556,'I',558-662,'E',664-711,'G',713-766 <DAR1>
A/Cross-references: UNIPARC:UPI0000052ACB; GB:M80536; NID:g181569; PIDN:AAA52308.1; PID:R;Gorvel, J. P.; Ferrero, A.; Chambraud, L.; Rigal, A.; Bonicel, J.; Maroux, S. Gastroenterology 101, 618-625, 1991
A/Title: Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human small intestine
A/Reference number: A61136; MUID:91317403; PMID:1677636
A/Accession: B61136
A/Molecule type: protein
A/Residues: 1-15,'X',17-22 <GOR>
A/Cross-references: UNIPARC:UPI0000172A2C

R.Boehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S. Biochem. J. 311, 835-843, 1995
A/Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a 1.5 kb promoter region
A/Reference number: S59510; MUID:96067939; PMID:7487939
A/Accession: S59510
A/Molecule type: DNA
A/Residues: 1-31 <BOB>
A/Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:g1195574; PIDN:AA835614.1; PID:R;Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg, J. Immunol. 149, 481-486, 1992
A/Title: Cloning and functional expression of the T cell activation antigen CD26.
A/Reference number: I56154; MUID:92325476; PMID:1352530
A/Accession: I56154
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-436,'S',438-766 <TAN>
A/Cross-references: UNIPARC:UPI000004F7BF; GB:M74777; NID:g180082; PIDN:AAA51943.1; PID:R;Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W. Immunogenetics 40, 331-338, 1994
A/Title: Genomic organization, exact localization, and tissue expression of the human CD26 gene
A/Reference number: S59857; MUID:95012454; PMID:7927537
A/Accession: S59857
A/Molecule type: DNA
A/Residues: 1-436,'S',438-766 <ABB>
A/Cross-references: UNIPARC:UPI000004F7BF; EMBL:U13734
C/Genetics:
A/Gene: GDB:DPP4
A/Cross-references: GDB:125239; OMIM:102720
A/Map position: 2q24.3-q24.3
A/Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3
C/Superfamily: dipeptidyl-peptidase IV
F/1-6/Domain: dipeptidylpeptidase hydrolase; glycoprotein; homodimer; proteinase; transmembrane
F/7-28/Domain: intracellular #status predicted <INT>
F/29-766/Domain: transmembrane #status predicted <TMN>
F/29-766/Domain: extracellular #status predicted <EXT>
F/85,92,150,219,229,281,321,520,685/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/630,708,740/Active site: Ser, Asp, His #status predicted
Alignment Scores:
Pred. No.: 2,97e-30 Length: 766
Score: 529.00 Matches: 186
Percent Similarity: 42.0% Conservative: 109
Best Local Similarity: 26.5% Mismatches: 280
Query Match: 9.5% Indels: 128
DB: 1 Gaps: 32
US-10-825-632-2 (1-3120) x CDHU26 (1-766)
QY 829 CCCGCTCATCCAGCTGGATT-----GCTTTTATACAT 861
Db 149 ProAsnAsnThrGlnTrpValThrTrpSerProValGlyHisLysLeuAlaTyrValTrp 168
QY 862 AGCAACCATATTTGGATATCTACATCGTACACGAGAGAAAGGAGACTCACTTATGTG 921
Db 169 AsnAsnAspIleTyrValIleGluProAsnLeuProSerTyrArgIleThrThr 188
QY 922 CACAATGAGTAGCAACATCGAAGAGAGTGCAGATCAGCTGGAGTCGCTTGTGT 981
Db 189 Gly-----LysGluAspIleIleTyrAsnGlyIleThrAspTrpVal 202
QY 982 CTCGAAGAAGAA---TTTGATAGATATCTGCTATTGGTGTGTCCTCAAAAGCTGAACA 1038
Db 203 TyrGluGluValPheSerAlaTyrSerAlaLeuTyrTrpSerProAsnGlyThrPhe 222
QY 1039 ACTCCAGTGGTGGTAAATTTCTAGATTCATATGAGAAATATGATGAATCTGAGGTG 1098
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluVal 233
QY 1099 GAAATATTTCAT-----GTTACATCCCTATGTTGGAAACAGAGGCGCAGATCA 1149
Db 234 ProLeuIleGluTyrSerPheTyrSerAspGluSerLeuGlnTyrProLysThrValArg 253
QY 1150 TTCGGTATCTCTAAACAGGTACAGCAAAATCTAAAGTCACCTTTTAAAGATGTCAGAAATA 1209

Db 254 ValProTyrProLysAlaGlyAlaValAenProThrValLysPhePheVal-----Val 271
 QY 1210 ATGATTCATCTCAAGAGGATCATAGATGTCATAGATAGAGAACTAAATCAACCTTTT 1269
 Db 272 AsnThrAspSerLeuSerValThrAenAlaThrSerIleGlnIleThrAlaProAla 291
 QY 1270 GAGATTCATTTGAAGGAGTTGAATATATGTCAGAGCTGGATGGACTCTCTGAGGAAAA 1329
 Db 292 SerMetLeuLeu-----GlyAspHisTyrLeuCysAspValThrTrpAla----- 306
 QY 1330 TATGCTGGTCCCTACTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCA 1389
 Db 307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu--- 316
 QY 1390 CCTGATTTATCCAGTAGAAGATGATGTTATGGAAGACAGAGACTCAATGAGTCA 1449
 Db 317 -----ArgArgIleGlnAen 321
 QY 1450 GTGCTGATTCGTGAGCCCACTAATATCTATGAAGAACACAGACATCTGGATAAAT 1509
 Db 322 Tyr-----SerValMetAspIleCysAspTyrAspLysSerSerClyArgTrp---Asn 338
 QY 1510 ATCCATGATCTTTTCATGTTTTTCCCAAGTCACAGAGAGAAATTCAGTTATTTT 1569
 Db 339 Cys-----LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrp 353
 QY 1570 GCCTCTGAATGCMAAACAGGTTTCCGTCAT-----TTATACAAA 1608
 Db 354 ValGlyArgPheArgProSerGluProHisPheThrLeuAspGlyAsnSerPheTyrLys 373
 QY 1609 ATTATCATCTTTTAAAGGAAGCAATATAAAGCATCCAGTGGTGGCTGCTGCTCCA 1668
 Db 374 IleIleSer-----AsnGluGluGlyTyrArgHisIle----- 384
 QY 1669 AGTGATTTCAAGTGCCTTCAAGAGGAGATAGCAATACCAAGTGGTGAATGGAGTT 1728
 Db 385 CysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGluVal 404
 QY 1729 CTGGCGCGCATGGATTAATATCAAGTTGATCAAGTCAAGAGCGCTGGTATATTTGAA 1788
 Db 405 IleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSer 419
 QY 1789 GGCACC---AAGACTCCCTTTAGAGCATCACCTGACGTAGTACGTAAATCTC 1845
 Db 420 AsnGluTyrLysGlyMetProGlyGlyArgAenLeuTyrLysIleGlnLeuIleAspTyr 439
 QY 1846 GAGAGGTGACAGGCTGACTACCGTGGCTACTCACATTTCTTCTGCATCAGT---CAG 1902
 Db 440 ThrLysValThrCysLeu-----SerCysGluLeuAsnProGlu 452
 QY 1903 CACTGTGACTCTTTTAAAGTAAAGTATAGTACCAAGAGAAATCCACAC-----TGT 1953
 Db 453 ArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCys 472
 QY 1954 -----GTGCTCTTTCAAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGT 2001
 Db 473 SerGlyProGlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArg 492
 QY 2002 ACAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGCTCTCTTCTGACTATCTCTCT 2061
 Db 493 ValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMet 509
 QY 2062 CCAGAA-----ATTTCCTCTTTGAAAGTACTACTGGATTTTACATTTGATGGATG 2112
 Db 510 ProSerLysLysLeuAspPheIleLeuAsnGluThrLysPhe---TrpTyrGlnMet 528
 QY 2113 CTCACAGGCTCATGATCTACAGCTGGAAAGAAATATCTACTGCTGCTGTCTATATAT 2172
 Db 529 IleLeuProProHis---PheAspLysSerLysLysTyrProLeuLeuAspValTyr 547
 QY 2173 GGTGCTCTCAGGTGCTGATGTTGTTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTG 2232

Db 548 AlaGlyProCysSerGln-----LysAlaAspThrValPheArgLeu 561
 QY 2233 AAT-----ACCCTAGCTCTCTAGGTATTATGGTGTGTAGTG---ATAGACACAGG 2280
 Db 562 AsnTrpAlaThrTyrLeuAlaSerThrGluAenIleIleValAlaSerPheAspGlyArg 581
 QY 2281 GGATCTCTCACCAGCGGCTTAAATTTGAAGCGCTTTAAATATAAATAATGGGTCAATA 2340
 Db 582 GlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAenArgArgLeuGlyThrPhe 601
 QY 2341 GAAATTCACCATCAGGTGGAAGGACCTCAATATCTAGCTTCTCGATATGATTTTCATGAC 2400
 Db 602 GluValGluAspGlnIleGluAlaAaGlnPhe---SerLysMetGlyPheValAsp 620
 QY 2401 TTAGATCGTGGGCATCCACGCTGCTCTATGAGGATACCTCTCCCTCATGCGCATTA 2460
 Db 621 AsnLysArgIleAlaIleTrpGlyTyrSerTyrGlyTyrValThrSerMetValLeu 640
 QY 2461 ATGCAGAGTTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCTCCTCTGTCGATC 2520
 Db 641 GlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGlu 660
 QY 2521 TTCATATGATACAGATACAGGACCTTATATGGT-----CACCTGACCAAGTAA 2574
 Db 661 TyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThrProGluAspAsnLeu 680
 QY 2575 CAGGCTATTATTAGATCTGGCCATCCAGCAGAGAAAGTTCCCTCTGAAACCAAT 2634
 Db 681 AspHisTyrArgAenSerThrValMetSerArgAlaGluAsnPhe-----LysGlnVal 698
 QY 2635 CGTTTACTGCTCTTACATGGTTCTCGATGAGATGATCCATTTTCACATACCATGATA 2694
 Db 699 GluTyrLeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGlnSerAlaGln 718
 QY 2695 TTACTGAGTTTTATGAGGCTGGAAGCATATGATTACAGATCTATCTCTCAGGAG 2754
 Db 719 IleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMetTrpTyrThrAspGlu 738
 QY 2755 AGACACAGCATTAAGATGTTCTGAAATCCGGGAGAACATTATGAATGCTCTTTGCTACT 2814
 Db 739 AspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThrHisMetSerHisPhe 758
 QY 2815 CTTCAAGAA 2823
 Db 759 IleLysGln 761
 RESULT 6
 T41703 dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, August 1998
 A:Reference number: Z22011
 A:Accession: T41703
 A:Status: preliminary; translated from GB/EMBL/DBDJB
 A:Molecule type: DNA
 A:Residues: 1-793 <MUR>
 A:Cross-references: UNIPROT:O14073; UNIPARC:UPI000013AB95; EMBL:AL031180; PIDN:CAA20138
 A:Experimental source: strain 972h-; cosmid c2E11 -chimeric
 C:Genetics:
 A:Gene: SPAC2E11.08
 A:Map position: 1
 C:Superfamily: dipeptidyl-peptidase IV
 Alignment Scores:
 Prod. No.: 1.46e-29 Length: 793
 Score: 519.50 Matches: 187
 Percent Similarity: 42.2% Conservatives: 113
 Best Local Similarity: 26.3% Mismatches: 256
 Query Match: 9.4% Indels: 155
 DB: 2 Gaps: 32

```
US-10-825-632-2 (1-3120) x T41703 (1-793)
QY 847 ATTGCTTTTATACATAGCAACCATATTTGGGATA-----TCTAACATCGTAACC 894
Db 183 LeuSerPheValTyrAsnAsnAspLeuTyrValArgLysAsnAspGlyAsnVal----- 200
QY 895 AGAGAAGAAGGAGCTACTAT-----GTGCACATAGCTAGCCAAC 939
Db 201 -----GlnArgLeuThrTyrAspGlyThrValAspValPheAsn----- 213
QY 940 ATGGAAGAAGATCCACATGACGTGAGTCGTACCTTTGTTCTCCAGAGAA-----TTT 996
Db 214 -----GlyLeuThrAspTrpIleTyrGluGluGluValLeu 225
QY 997 GATGATATTTCTGGCTATTGGTGTCTCAAAAGCTGAAACAACTCCACGTGGGTGATAA 1056
Db 226 SerSerProSerThrIleTrpSerProAspSerAsp----- 238
QY 1057 ATTCTTTGAAATTTATATGAGAAGAAATGATGAATCTGAGGTGGAAATTAATTCATGTT--- 1113
Db 239 -----LysIleAlaPheLeuLysLeuAsnGluSerGluIleProThrTyrHisTyrPro 256
QY 1114 -----ACATCCCTATGTTGGAAACACAGGCGGAGATTC 1149
Db 257 LeuTyrThrAlaGluLeuAspProSerLeuProGluPheAspTyrAsnLysAspMetAla 276
QY 1150 TTCGGTTATCTCTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTCAGAAATA 1209
Db 277 IleLysTyrProLysProGlyAsnProAsnProSerValSerLeuPheValAlaAspLeu 296
QY 1210 ATGATTGTAGCTGAAGAGGATCATAGATGTCATAGATAGGAACCTAATTCACCTTTT 1269
Db 297 AsnSerAsnAlaSerSerAsnPheSerLeuTrpHisAsnGluProLeuAlaGluPro--- 315
QY 1270 GAGATTCTATTGAAGAGCTTGAATATATTGCGAGAGCTGATGGAATCTCTGAGGAA 1329
Db 316 -----ValValGlnAsnValLeuTrpValAsnThrSer----- 326
QY 1330 TATGCTTGCTCCATCTACTA-----GATCGCTCCAGACTCCGCTACAGATAGTG 1380
Db 327 -----SerValLeuValGlnPheThrAsnArgAsnSerThrCysIleThrAlaArg 343
QY 1381 TTGATCTCACCTGAATTA-----TTTATCCAGTAGAAGATGATGTTATGGAAGG--- 1431
Db 344 LeuLeuAspThrGluLeuLysSerIleHisThrValLysThrGluCysLeuGluGluGly 363
QY 1432 -----CAGAGACTCATGAGTCAGTCGCTGATTCGTGACGCCCACTAATTATC 1479
Db 364 TrpTyrGluValGlnSerAlaLysMetPhePro-----LeuAsnAsnSerLeuVal 381
QY 1480 TATGAAGAACAACAGACATCTGGATAATATCCATGACATCTTTCATGTTTTCGCCAA 1539
Db 382 TrpGluAsnTrpSerAsp-----GlyTyrPheAspIleLeuAlaLeuAspAspTyr 398
QY 1540 AGTCAGGAAGAGGAATTCAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCAT 1599
Db 399 AsnHisLeuAlaPheIleProPhe----- 406
QY 1600 TTATACAAAATATCATCTATTTTAAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTG 1659
Db 407 -----AsnGlySerSer 410
QY 1660 CTGCTCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACAGTGGTAA 1719
Db 411 Pro-----IleTyrLeuThrSerGlyAla 418
QY 1720 TGGGAAGTCTTGGCGGATGATCTAATATCAAGTTGATGAAGTCAGAGCGCTGGTA 1779
Db 419 TrpAspValThr-----AspGlyProIleHisIleAspGlyAspPheGlyAsnVal 435
QY 1780 TATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTAGCTAGTACGTACGTACGTA 1839
Db 771 SerGluTrpIleGlyAspAlaLeuGlyArgIle 781
```

```
Db 436 TyrPheLeuAlaThrLeuLysAspSerThrGluArgHisLeuTyrTyrValSer---Leu 454
QY 1840 AATCTGGAGAGGTGACAAGGCTGACTGACGCTGGCTACTCA-----CATTTGTC 1890
Db 455 AspThrLeuGluIleTyrGlyIleThrAspAsnGlyGluAspGluGlyTyrTyrSerThr 474
QY 1891 TGCATCAGTCAGCACTGCTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACAC 1950
Db 475 SerPheSerProPheGlyAspPheTyrValLeuAsnTyrHisGlyProAspValPro--- 493
QY 1951 TGTGTGTCCTTTTACAAAGCTATCAAGTCTCAAGATGACCCCAACTTGTCAAAACAAAGAA 2010
Db 494 -----TrpGlnGluLeuArgSerThrLysAspLysAspTyrCysLeuSerLeuGlu 510
QY 2011 TTTTGGGCCCACTTTTGGATTTCAGCAGGTCTCTCTCTCTGACTATATCTCTCCCAAAAT 2070
Db 511 ThrAsnSerArgLeuLysGlnGln-----LeuSerSerIleThrLeuProSerVal 527
QY 2071 -----TTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTAC 2118
Db 528 GluTyrGlyLysLeuThrPheAsnAspThrThr---PheAsnPhe-----MetGluArg 544
QY 2119 AAGCCTCATGATCTACAGCCTCGAAGAAATATCTCTCTCTCTCTATATATGCTGCT 2178
Db 545 ArgProArgAsnPheAspValAsnLysLysTyrProValLeuPhePheAlaTyrGlyGly 564
QY 2179 CCTCAGGTGAGTGTGGTGAATAATCGTTTAAAGAGGACCAAGTATTTCCGCTTGAATACC 2238
Db 565 ProGlySerGlnGlnVal-----AlaLysLeuPheArgValAspPhe 578
QY 2239 CTAGCTCTCTAGGT-----TATGCTGTTGTAGTATGATGACAAACAGG 2280
Db 579 GlnAlaTyrLeuAlaSerHisProAspPheGluPheIleValValThrLeuAspGlyArg 598
QY 2281 GGATCTCTGTCACGCGGCTTAAATTTGAAGCGGCTTTTAAATATAAATGGTCAATA 2340
Db 599 GlyThrGlyPheAsnGlyAsnAlaPheArgTyrSerValSerArgHisLeuGlyGluTrp 618
QY 2341 GAAATTCACCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCCGATATGATTTCATTGAC 2400
Db 619 GluSerTyrAspGlnGlyGlnAlaGlyLysPheTrpAlaAsp---LeuProPheValAsp 637
QY 2401 TTAGATCGTGTGGCATCCACGCTGCTCTATGAGGATACCTCTCTCTGATGCAATTA 2460
Db 638 GluAsnHisValGlyIleTrpGlyTrpSerTyrGlyTyrLeuThrLeuLysThrLeu 657
QY 2461 ATGACAGGTCACATATCTTCAGGTTGCTATGCTGGGCGCCAGTCACTCTGTGGATC 2520
Db 658 ---GluThrGlnAspValPheSerTyrGlyMetAlaValAlaProValThrAspTrpArg 676
QY 2521 TTCTATGATACAGGATACACGGAACGTTATATGCGTCCCTGACCAAGATGAACAGGCG 2580
Db 677 LeuTyrAspSerValTyrThrGluArgTyrMetAspLeuProGlnTyrAsnLysGluGly 696
QY 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCGTTTA 2640
Db 697 ---TyrLysAsnSerGlnIleHisAspTyrGluLysPhe---LysGlnLeuLysArgPhe 714
QY 2641 CTGCTCTTACATGCTTCTCGGATGAGATGCAATGTCATTTTTCACATACCATGATATTACTG 2700
Db 715 PheValAlaHisGlyThrGlyAspAsnValHisPheGlnHisSerMetHisLeuMet 734
QY 2701 AGTTTTTTAGTGGGCTGGA---AAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGA 2757
Db 735 AspGlyLeuAsnLeuAlaAsnCysTyrAsnTyrAspMetAlaValPheProAspSerAla 754
QY 2758 CACAGAAACCTTGATTCCTGAATCGGAGAACATTATGAATCACTGCACTTTTTCACCTT 2817
Db 755 HisSerIle-----SerTyrHisAsnAlaSerLeuSerIleTyrHisArgLeu 770
QY 2818 CAGAAACCTTGATCAGTATTGCTGCTCTA 2850
Db 771 SerGluTrpIleGlyAspAlaLeuGlyArgIle 781
```

RESULT 7
 S23752 dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
 N:Alternate names: CD26 alpha subunit; THEM alpha subunit
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S23752; A46465; A56030
 R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
 J. Biol. Chem. 267, 2200-2208, 1992
 A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di-
 A:Reference number: S23752; MUID:92129288; PMID:1370813
 A:Accession: S23752
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-760 <MAR>
 A:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X59384
 R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
 J. Immunol. 147, 447-454, 1991
 A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase
 A:Reference number: A46465; MUID:91302787; PMID:1712807
 A:Accession: A46465
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-20 <VIV>
 A:Cross-references: UNIPARC:UPI0000172A2E
 A:Experimental source: M14 T thymoma cells, Swiss nu/nu
 A:Note: sequence extracted from NCBI backbone (NCBI:P:42236)
 R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
 Biochemistry 33, 15204-15214, 1994
 A:Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.
 A:Reference number: A56030; MUID:95092780; PMID:7999781
 A:Accession: A56030
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 746-760 <BBR>
 A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620
 C:Genetics:
 C:Superfamily: dipeptidyl-peptidase IV
 C:Keywords: dipeptidyl-peptide hydrolase; glycoprotein; transmembrane protein
 F:213,223,315,514,679/Binding site: carbohydrate (Asn) (covalent) #status Predicted
 F:624,702,734/Active site: Ser, Asp, His #status Predicted

Alignment Scores:
 Pred. NO.: 1.71e-29 Length: 760
 Score: 518.50 Matches: 203
 Percent Similarity: 38.1% Conservative: 107
 Best Local Similarity: 25.0% Mismatches: 306
 Query Match: 9.3% Indels: 197
 DB: 1 Gaps: 37

US-10-825-632-2 (1-3120) x S23752 (1-760)

QY	604	ATGTATCTCGAGAGAGAACTATTATAGAGAAAGAAACGCATTGGAAACAGTCGGAATT	663
Db	67	LeuTyrLysGlnGluAsnAsnLeuLeuLeuLeuAsnAlaGluHisGlyAsnSerSerIle	86
QY	664	-----GCTCTTACGATTATCACCAGGAAGTGGAAACATTCTCG	702
Db	87	PheLeuGluAsnSerThrPheGluSerPheGlyTyrHisSerValSerProAspArgLeu	106
QY	703	TTT-----CAAGCCGGTAGT	717
Db	107	PheValLeuLeuGluTyrAsnTyrValLysGlnTrpArgHisSerTyrThrAlaSerTyr	126
QY	718	GGAAATTTATCAGCTAAAGATGAGGGCCCAAGGATTTACGCAACACCTTTAAGGCC	777
Db	127	AsnIleTyrAspValAsnLys-----ArgGln	135
QY	778	AATCTAGTGGAAACTAGTTGTCCTCCCAACATACCGATGGATCCAAATATATGCCCGCTGAT	837
Db	136	LeuIleThrGluGluLysIleProAsn-----Asn	145

QY	838	CGAGACTGGATT-----GCTTTATATACATAGCAACGAT	870
Db	146	ThrGlnTrpIleThrTrpSerProGluGlyHisLysLeuAlaTyrValTrpLysAsnAsp	165
QY	871	ATTGGATATTAACATCGTAACACAGAGAAGAGAGACTCACTTATGTGCACATGAG	930
Db	166	IleTyrValLysValGluProHisLeuProSerHisArgIleThr-----	180
QY	931	CTAGCCACATCGGAAGAGATGCACATCAGCTGGAGTCGCTACCTTTCTCCAGAA	990
Db	181	---SerThrGlyGluGluAsnValIleTyrAsnGlyIleThrAspTrpValTyrGluGlu	199
QY	991	GA-----TTTGATAGATATTCCTGGCTATTGGTGGTGCTCCAAA	1029
Db	200	GluValPheGlyAlaTyrSerAlaLeuTrpTrpSerProAsnAsnThrPheLeuAlaTyr	219
QY	1030	GCTGAAACAACTCCAGCTGGGTAAATCTTTAGAAATCTATATGAAGAAATATGATGA	1089
Db	220	AlaGlnPheAsnAspThrGlyValProLeuIleGluTyrSerPhe---TyrSerAspGlu	238
QY	1090	TCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACAAGAGGGCAGATTCA	1149
Db	239	Ser-----LeuGlnTyrProLysThrValTrp	247
QY	1150	TTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGCAGAAAT	1209
Db	248	IleProTyrProLysAlaGlyAlaValAsnProThrValLysPhePheIle-----Val	265
QY	1210	ATGATTGATGCTGAAGGAGGATCATAGATGTCATAGTAAGGAACATAATCAACCTTTT	1269
Db	266	AsnIleAspSerLeuSerSerSerSerSerAlaProIleGlnIleProAlaProAla	285
QY	1270	GAGATTCTATTTGAAGGAGTTGAATATATTCGCAGAGCTGGATCGACTCTCGAGGAAA	1329
Db	286	SerValAla---ArgGlyAspHisTyrLeuCysAspValValTrpAlaThrGluGluArg	304
QY	1330	TATGCT-----TGGTCCATCTCTACTAGTAGTCGCTCCAG	1362
Db	305	IleSerLeuGlnTrpLeuArgIleGlnAsnTyrSerValMetAlaIleCysAspTyr	324
QY	1363	ACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTT	1422
Db	325	AspLysIleAsnLeuThrTrpAsnCysProSer-----	335
QY	1423	ATGMAAGCAGACACTCATTCAGTCAGTCGCTGATCTGTGACGCCACTAATTATCTAT	1482
Db	336	---GluGlnGlnHisVal-----	340
QY	1483	GAAGAAACACACACATCTGGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGT	1542
Db	341	GluMetSerThrThrGlyTrpValGly-----ArgPheArgProAlaGluProHisPhe	358
QY	1543	CACGAGAGAAATTGAGTTTATTTTGGCTCTGAATGCAGAAACAGGTTTCCGTCATTTA	1602
Db	359	ThrSerAspGlySerSerPheTyrLysIleIleSerAspLysAspGlyTyrLysHisIle	378
QY	1603	TACAAATTCATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGTGGCTGCCT	1662
Db	379	CysHisPhe-----	381
QY	1663	GCTCCAGTAGTATTTCAG-----TGTCTATCAAGAGGAGATAGCAATACCACTGGTGA	1719
Db	382	---ProLysAspLysLysAspCysThrPhe-----IleThrLysGlyAla	395
QY	1720	TGGGAAGTTCTTGCCCGCATGGATCTAATATCAAGTTGATGAAGTCAAGAGGCTGTA	1779
Db	396	TrpGluValIle-----SerIleGluAlaLeuThrSerAspTyrLeuTyr	410
QY	1780	TATTTTGAAGCCACC-----AAAGACTCCCTTTAGAGCATCACCTGCTAGTCTAGTTAC	1836
Db	411	TyrIleSerAsnGlnTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIle-----	428

Qy 1672 GATTCAAGTGTCTATCAAGAGGAGATAGCAATTACCACTGGTGAATGGGAAGTTCTT 1731

Qy 1732 GSCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGCGCTGATATATTTTGAAGGC 1791
Db 407 -----SctIleGluAlaLeuThrSerAspTyrLeuTyr-Tyr-IleSerAsn 421
Qy 1792 ACC---AAAGACTCCCTTTAGAGCATCACCTGTACCTGTACCTAGTACGTTAACTCCTGGA 1848

100

[illegible]

QY 2683 CATACGATATATTTACTGAGTTTCTTGTAGTGGCTGGAAGCAATATGATTTACAGATC 2742
Db : : : : :
716 GlnSerAlaGlnIleSerLysAlaLeuValAspAlaGlyValAspPheGlnAlaMetTrp 735
QY 2743 TATCCTCAGGAGACACAGCATAGAGTTCTCTCAATCGGAGAACATTTAGTACGTCAT 2802
Db : : : : :
736 TyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrSerHis 755
QY 2803 CTTTTCACCTACCTTCAAGAA 2823
Db : : : : :
756 MetSerHisPheLeuGlnGln 762

RESULT 9
A30107
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YHR028c
C;Species: Saccharomyces cerevisiae
C;Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C;Accession: S46780; A30107
R;Du, Z.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of S. cerevisiae cosmid 8082.
A;Reference number: S46773
A;Accession: S46780
A;Molecule type: DNA
A;Residues: 1-818 <DUZ>
A;Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031A5F; EMBL:U10399; NID:g500689; PID
R;Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an in
A;Reference number: A30107; MUID:89174971; PMID:2647766
A;Accession: A30107
A;Molecule type: DNA
A;Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRRLET', 199-199, 'N', 201-365, 'DFKRGKERF', 376-57
A;Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A;Note: the authors translated the codon ACC for residue 572 as Asn
C;Genetics:
A;Gene: SGD:DAP2; STE13; MIPS:YHR028c
A;Cross-references: SGD:S0001070, MIPS:YHR028c
A;Map position: 8R
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacu
F;30-45/Domain: transmembrane #status predicted <TM>
F;63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 5.08e-27 Length: 818
Score: 484.50 Matches: 183
Percent Similarity: 39.4% Conservative: 95
Best Local Similarity: 26.0% Mismatches: 250
Query Match: 8.7% Indels: 177
DB: 1 Gaps: 35

US-10-825-632-2 (1-3120) x A30107 (1-818)

QY 847 ATTGCTTTTATACATAGCAACATATTTGGATATCTAACATCGTAACACGAGAGAAAGG 906
Db : : : : :
201 IleAlaTyrValGlnAspAsnAsnIleTyrIleTyrSerAlaIleSerLysThrIle 220
QY 907 AGACTCACTTATGTGCACAAATGAG-----CTAGCCAACTGGGAAGAGATGCC 954
Db : : : : :
221 ArgAla-----ValThrAsnAspGlySerSerPheLeuPheAsnGlyLysProAsp--- 237
QY 955 AGATCAGCTGGAGTCGCTACCTTGTCTTCCCAAGAGAA-----TTTGATAGATATTTCTGGC 1011
Db : : : : :
238 -----TrpValTyrGluGluValPheAspAspLysAla 250
QY 1012 TATTGGTGGTGTCCAAAGCTGAACAACTCCAGTGGTGTAATAATCTTAGAATCTTA 1071
Db : : : : :
251 AlaTrpTrpSerProThrGlyAsp-----TyrLeuAla 261
QY 1072 TATGAAGAAATGATGAATCTCAGGTG---GAATATTATTCATGTTACATCCCTATGTTG 1128
Db : : : : :
1128

Db 262 PheLeuLysIleAspGluSerGluValGlyGluPheIle-----IleProTyrTyr 278
QY 1129 GAAACAAGGAGGCGAGAT-----TCATCCGTTATCTTAAACACAGT 1170
Db : : : : :
279 ValGlnAspGluLysAspIleTyrProGluMetArgSerIleLysTyrProLysSerGly 298
QY 1171 ACAGCAAAATCCTTAAAGTCACTTTTAAGATGTCAGAAAATAGTATGATCTCGAAGGAAGG 1230
Db : : : : :
299 ThrProAsnPro-----HisAlaGluLeuTrp 307
QY 1231 ATCATAGATGTCATAGTAAGGAACATAATTCAACCT-----TTTGAGATTTCTATTGAA 1284
Db : : : : :
308 ValTyrSerMetLysAspGlyThrSerPheHisProArgIleSerGlyAsnLysAsp 327
QY 1285 GGAGTTGAATATATATCCAGAGCTGGATGGACTCTCGAGGAGAAAATATGCTTGGTCCATC 1344
Db : : : : :
328 GlySerLeuLeuIleThrGluValThrTrpValGlyAsnGly-----AsnVal 343
QY 1345 CTACTA-----GATCGTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAA 1395
Db : : : : :
344 LeuValLysThrThrAspArgSerSerAspIleLeuThrValPheLeuIle----- 360
QY 1396 TTATTTATCCAGTAGAAGATGCTTTATGGAAAGCAGAGACTCATTTGACTCAGTGCT 1455
Db : : : : :
361 -----AspThrIleAlaLysThrSerAsnValValArgAsn----- 372
QY 1456 GATTCTGTGAGCGCCACTAATTATCTATGAAGAAACACAGACATCTGTGATAAATATC--- 1512
Db : : : : :
373 -----GluSerSerAsnGlyGlyTrpTrpGluIleThr 383
QY 1513 CATGACATCTTTCATGTT-----TTTCCCAAGTCAAGAGAGAAATTT 1557
Db : : : : :
384 HisAsnThrLeuPheIleProAlaAsnGluThrPheAspArgProHisAsnGlyTyrVal 403
QY 1558 GAGTTTATTTTCCCTCTGAATGCAAAACAGGTTTCCGCTATTATATACAAATTACATCT 1617
Db : : : : :
404 AspIleLeuProIleGly-----GlyTyrAsnHisLeu----- 414
QY 1618 ATTTTAAAGGAACAAATATAACGATCCAGTGGTGGCTGCCTGCTCCAAGTGATTTTC 1677
Db : : : : :
415 -----AlaTyrPheGluAsnSerAsnSer-----SerHisTyr 425
QY 1678 AAGTGTCTTATCAAGAGGAGATAGCAATTTACAGTGTGTAATGGGAATGTTCTGGCGG 1737
Db : : : : :
426 Lys-----ThrLeuThrGluGlyLysTrpGluValVal----- 436
QY 1738 CATGGATCTAATATCCAAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1797
Db : : : : :
437 ---AsnGlyProLeuAlaPheAspSerMetGluAsnArgLeuTyrPheIleSerThrArg 455
QY 1798 GACTCCCTTTAGAGCATCACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1857
Db : : : : :
456 LysSerSerThrGluArgHisValTyrTyrIleAspLeuArgSerProAsnGluIleIle 475
QY 1858 AGGCTGACTCACGCTGCTACTCA-----CATTCTTGTGTCATGATGATGATGATGATGATGAT 1911
Db : : : : :
476 GluValThrAspThrSerGluAspGlyValTyrAspValSerPheSerSerGlyArgArg 495
QY 1912 TTCTTTATAAGTAGTATAGTACCAAGAGAGAAATCCACAC----- 1950
Db : : : : :
496 PheGlyLeuLeuThrTyrLysGlyProLysValProTyrGlnLysIleValAspPheHis 515
QY 1951 -----TGT-----GTGTCCTTTTACAAGCTA 1971
Db : : : : :
516 SerArgLysAlaGluLysCysAspLysGlyAsnValLeuGlyLysSerLeuTyrHisLeu 535
QY 1972 TCAAGTCTCAAGATGACCAACTTGCACAAAACAAAGAAATTTTGGGCCCACTTTTGGAT 2031
Db : : : : :
536 GluLysAsnGlu-----ValLeuThr 542
QY 2032 TCAGCAGGTCTCTTCTGCTGACTATACCTCTCCAGAAATTTTCTCTTTGGAAGTACTACT 2091
Db : : : : :
543 LysIle-----LeuGluAspTyrAlaValProArgLysSerPheArgGluLeuAsnLeu 560

QY 2092 GGATTTCATATTGTATGG-----ATGCTCTACAAGCCTCATGATCTACAGCCT----- 2139
Db 561 GlyAspGluPheGlyLysAspIleLeuValAsnSerTyrGluIleLeuProAsnAsp 580
QY 2140 -----GGAAAGAAATATCTTACTGTGCTGTTCATATATATGCTGCTCTCAG 2184
Db 581 PheAspGluThrLeuSerAspHisTyrProValPhePheAlaTyrGlyGlyProAsn 600
QY 2185 GTGCAAGTGGTGAATAATCGTTAAAGGAGTCAAGTATTTCCGCTTGAAT----- 2235
Db 601 SerGlnGlnVal-----VallyeThrPheSerValGlyPheAsnGlu 614
QY 2236 ---ACCTAGCCTCTAGTATGCTGTGATGATAGACCAACAGGGGATCCTGTCAC 2292
Db 615 ValValAlaSerGlnLeuAsnAlaIleValValValAspGlyArgGlyThrGlyPhe 634
QY 2293 CGAGGGCTAAATTTGAAGGCGCTTAAATATAAATGGGTCAATAGAAATGACGAT 2352
Db 635 LysGlyGlnAspPheArgSerLeuValArgAspArgLeuGlyAspTyrGluAlaArgAsp 654
QY 2353 CAGGTGGAGGACTCCAATATCTAGCTTCGATATGATTTCAATGACTTAGATCGTGTG 2412
Db 655 GlnIleSerAlaAlaSerLeuTyrGlySer---LeuThrPheValAspProGlnLysIle 673
QY 2413 GGCATCCAGCGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCA 2472
Db 674 SerLeuPheGlyTrpSerTyrGlyGlyThrLeuThrLeuLysThrLeuGlyLysAspGly 693
QY 2473 GAT---ATCTTCAGGGTGTCTATTGCTGGGCGCCAGTCACCTCTGTGGATCTTCTATGAT 2529
Db 694 GlyArgPhePheLysTyrGlyMetSerValAlaProValThrAspTrpArgPheTyrAsp 713
QY 2530 ACAGGATACACGGAAGCTTATATGGTCACCCCTGACAGAAATGAACAGGGCTATTACTTA 2589
Db 714 SerValTyrThrGluArgTyrMethIsthrProGlnGluAsnPheAspGlyTyrValGlu 733
QY 2590 GGNATCTGTG-----GCCATGCAACGACAGAAAGTTCCTCCTGCAACCAATCTG 2637
Db 734 SerSerValHisAsnValThrAlaLeu-----AlaGlnAlaAsnArg 747
QY 2638 TTAATCTCTTACATGCTTCTGATGATGAGATGCCATTTTGCACATACCATATATTA 2697
Db 748 PheLeuLeuMetHisGlyThrGlyAspAspAsnValHisPheGlnAsnSerLeuLysPhe 767
QY 2698 CTGAGTCTTTTATGAGGGCTGGA---AAGCCATATGATTTACAGATCTATCTCCTCAGGAG 2754
Db 768 LeuAspLeuLeuAspLeuAsnGlyValGluAsnTyrAspValHisValPheProAspSer 787
QY 2755 AGACACAGCATAGA 2769
Db 788 AspHisSerIleArg 792

RESULT 10
A49737
dipeptidyl aminopeptidase (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O5045; protein YOR219c; protein YOR50-9
C;Species: Saccharomyces cerevisiae
C;Date: 03-Mar-1994 #sequence revision 03-Mar-1994 #text_change 09-Jul-2004
C;Accession: A49737; S45451; S60946; S67112; S71721
R;Santa Anna-A, S.; Herskowitz, I.
submitted to the Protein Sequence Database, July 1993
A;Reference number: A49737
A;Accession: A49737
A;Molecule type: DNA
A;Residues: 1-931 <SN>
A;Cross-references: UNIPROT:P33894; UNIPARC:UPI0000136060; GB:L21944; NID:G347196; PIDN:
Y;Anna-Arriola, S.; Herskowitz, I.
Yeast 10, 801-810, 1994
A;Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase
A;Reference number: S45451; MUID:95066382; PMID:7975897
A;Accession: S45451
A;Molecule type: DNA

A;Residues: 1-931 <ANN>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:G347196; PIDN:AAA35119.1; PI
R;Gallison, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A;Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
A;Reference number: S60938
A;Accession: S60946
A;Molecule type: DNA
A;Residues: 1-931 <GAL>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:G1050762; PIDN:CAA63182.1;
R;Boyer, J.; Fairhead, C.; Gallion, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67104
A;Accession: S67112
A;Molecule type: DNA
A;Residues: 1-931 <BOY>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:Z75127; NID:G1420507; PIDN:CAA99437.1;
R;Gallison, F.; Dujon, B.
Yeast 12, 877-885, 1996
A;Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV
A;Reference number: S71721; MUID:96437977; PMID:8840505
A;Accession: S71721
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-931 <GAW>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:G1050762; PIDN:CAA63182.1;
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:STE13; YCII
A;Cross-references: SGD:S0005745; MIPS:YOR219c
A;Map position: 15R
C;Function:
A;Description: involved in processing of alpha-factor prepropheromone
C;Keywords: dipeptidylpeptidase hydrolase; glycoprotein; transmembrane protein
F;113-150/Domain: transmembrane #status: predicted <TM>
F;377/Binding site: carboxylate (Asn) (covalent) #status predicted
F;785,863,896/Active site: Ser, Asp, His #status predicted

Alignment Scores:
Pred. No.: 1,36e-25 Length: 931
Score: 465.00 Matches: 218
Percent Similarity: 36.2% Conservative: 150
Best Local Similarity: 21.5% Mismatches: 380
Query Match: 8.4% Indels: 268
DB: 2 Gaps: 41

US-10-825-632-2 (1-3120) x A49737 (1-931)

QY 229 GAACAGACAGCTGGTGTGAGATATTGAACTGGGACTGTGAGAGAAATATTGAA 288
Db 43 GlnSerAsnGluAsnGlyHisThrIleAsnGluIleArgProThrGluAlaThrIleAsp 62
QY 289 TCACAGATCGCGCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTT 348
Db 63 ValThrAspValProGln---ThrProPheLeuGlnGluGlnTyrSer----- 77
QY 349 AAAAAGCTGCTTCCGATACCAAGAAATATCATGCTACATGCTAAGGCACCATCAT 408
Db 78 -----MetArgProArgArgGlu 83
QY 409 GATTTCATGTTTGTGAAGAGATGATCCAGATGACCTTCATTCAGACAGATCTATTAC 468
Db 84 SerPheGlnPhe-----AsnAspIleGluAsnGlnHisThrHisSerPhePhe 100
QY 469 CTTCCTCATGTCT-----GGTGCAGACAGAGAAAT 498
Db 101 SerValAsnLysPheAsnArgGlyTrpGlyGluTrpSerLeuProGlnLysArgSerTyr 120
QY 499 ACACCTGTTTATTTCTGAAATTCACAAATCCAAATATGAGACGAGCTCTTAATGCTCTCT 558
Db 121 ValLeuValPheThrLeuIleAlaLeuSerValLeuValLeuValLeuIlePro 140

Tue Apr 18 08:18:36 2006

550 ATGCTCTCTTGGGAAGCCTCTTTGG-----GATCCTTTT 582
153 -----AenProLeuLeuPheGluValGluLeuPheArgGluLeuSer 168
583 CAGGCAACACTGGGACTATGAAATGTATCTTCGAGAAGAAGAACTATTAAAGAGAAA 642
169 AspSerArgIleThrTyraSpIleGlyLeuArgLysGluLeuVal----- 184
643 CGCATTTGAACAGTCGGAAATGCTCTTCTACGATTATCACCAGGAAGTGGAACTTTCTG 702
185 -----IleGlnAlaPheLysTrpAsnGlyLysPheAsnAspPheVal 198
703 TTTCAAGCCGGTAGTGGAAATTTATCATCGTAAAGATGGAGGCCACAAGGATTTCGCAA 762
199 PheValGluSerAsnAlaValIleTyr----- 207
763 CAACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCTCCAACATACGGATGGATCCAAA 822
208 -----GlnSerSerPro----- 211
823 TTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTTCGATATCT 882
211 ----- 211
883 AACATCGTAACACAGAGAAGAAGAGACTCATTTATGTGCAAAATGAGTAGCCCAACATG 942
212 -----GluGluGluGlyLeuThrArgValSerAsnGlyGlyGluHisThr 226
943 GAAGAGATGCCAGATCAGCTGGAGTCCCTACCTTTGTCTTCCAAAGAAGAA-----TTTGAT 999
227 ValAsp-----GlyLeuPheAspTrpIleTyrGluGluGluIlePheGly 241
1000 AGATATCTTGGCTATTGGTGGTGTCAAAAGCTGAAACAACTCCACAGTGGTGGTAAATTT 1059
242 ArgLysAspAlaMetTrpTrpSerThrLys-----GlyAspGlnLeu 255
1060 CTTAGAATCTTATGCAAGAAATGATGATCTGAGGTGGAAATATTATTCATGTTACATGCC 1119
256 AlaTyrAlaSerTyrAspAsnHisLeuThrLysAsnValSerLeu-----LysThrTyr 273
1120 CCTATGTTGGAAACAAGGAGGGCAGAT-----TCATTCCGTTATCTCTAAACA----- 1167
274 HisArgLeuGluProTyrProIleAspThrAsnPheHisTyrProLysThrPheAlaLys 293
1168 -----GGTACAGCAATCCTAAAGTCATCTTTAGATGTCACGAATAATGATTGATGCT 1221
294 ValLeuProThrTyrThrLeuSerIleTrpAsnLysLysThrGlu----- 308
1222 GAAGCAAGGATCATAGATGTTCATAGATAGAACTAATTCACACTTTTCAGATTCTA----- 1278
309 GlnSerArgGlnLeuAspValGlnLeuLysAspSerLeuSerTyrHisTyrLeuLeuAla 328
1279 -----TTTGAAGCAGATTGAAATATATATATCCAGAGCTGGATGGACTCTCTGAG 1323
329 ValLysTrpLeuGluIleAsnGlyThrGluGlnLeuValSerVal-----TrpThr----- 345
1324 GGAAATATGCTTGGTCCATCCTACTA-----GATCGCTCCACAGACTCCG 1368
346 AsnArgTyrGlnAsnGluValAlaLeuThrIleCysAspTrpAspThrAlaIleCysArg 365
1369 CTACAGATAGTGTGATCTCACCTGAATTATTTATCCAGTAGAAGATGATGTTATGGAA 1428
366 LeuGluPheGlu----- 369
1429 AGGCAGAGACTCATTTGACTGAGTCGCTGATTCGTGAGCCCACTAATTATCTATGAGAA 1488
370 -----TyrLysTyr 372
1489 ACAACAGACATCTGGATAAATATCCATGACATCTTTGATGTTTCCCAAGACTCAGCAA 1548
373 AlaSerLysArgTrpValThr-----HisAspAspPheHisSerIle-----ThrSerPheGlu 390
1549 GAGGAAATTCAGTTATTTTTCCTCTGAATGGCAAAACAGGTTTCCGTCATTTATACAA 1608

QY	1324	GGAAATATGCTTGGTCACCTCACIA-----					
Db	346	AenArgTyrGlnAsnGluValAlaLeuThrIleCysAspTrpAspThrAlaIleCysArg	::::	::::	::::	::::	::::
QY	1369	CTACAGATAGTGTTGATCTCACTCGAATTATTTATCCCATAGAGAATGATTGTGGAA					
Db	366	LeuGluPheGlu-----					
QY	1429	AGGCAGACTCATTTGATCGCTGATTTCTGTGAGCCACTAAATTATCTATGAAGAA					
Db	370	-----	----	----	----	TyrLysTyr	372
QY	1489	ACAACAGACATCTGATAAATATCCATGACATCTTTCATGTTTTCCCACAAAGTCACGAA					1548
Db	373	AlaSerLysArgTrpValThr---HisAspAspPheHisSerIle---ThrSerPheGlu					390
QY	1549	GAGGAAATGTAGTTTATTTTGGCTCTCTGAATGCGAAAACAGGTTTCCGTCATTTATACAAA					1608

US-10-825-632-2 (1-3120) x T25174 (1-799)

[illegible]

```
Db 391 AspThrLeuPhePheLeuLeuPro-----HisAspLysArg 402
Qy 1609 ATTACATCTATTATTAAGGAAGCAAAATATATAACGATCCAGTGGTGGCTGCTCTCA 1668
Db 403 AspAsnAlaPheGlnGlnValAlaSerLeuArgLeuSerHisGlyGlnLeuArgThrPro 422
Qy 1669 AGTGATTTCAAGTGTCTTATCAAGAGGAGATAGCAATATACAGTGGTGAATGGGAAGTT 1728
Db 423 Lys-----PheLeuAsnLeuGluTyArgVal 432
Qy 1729 CTGTGGCGCATGATCTAATATCCAGTTGATGATCAGAGGCTGGTATATTGAA 1788
Db 433 ThrSerIleAsnGlyIleAsnLysGluThr-----ArgThrIlePhePheHis 448
Qy 1789 GGCACCAAGATCCCTCTTAGAGCATCACCTGTACGTAGTACGTAGTAAATCCTGGA 1848
Db 449 AlaAlaAlaProLysProSerHisArgSerLeuPheSerTySer----- 463
Qy 1849 GAGGTGACAGCTGACTGACCGTGGTACTCATCTTGTGCTGATCAGT----- 1899
Db 464 -----LeuAlaAspGluSerArgAsnSerAlaTyrcysIleSerCysSerIle 479
Qy 1900 CAGCACTGTGACTTCTTTATAGTAAGTATAGTAAACCAGAGAATCCACACTGTGTGCTC 1959
Db 480 LysAsnCysThrThrPalaGlnAlaGlnMetAspAspGlnMetLysThrAlaIleValSer 499
Qy 1960 CTTTCAAGCTCAAGTCCCTGAAGAT-----GACCCAACCT 1995
Db 500 CysLysGlyProAlaAlaProHisThrAlaIleValAsnLeuThrArgMetAspSerAsp 519
Qy 1996 TGCAAAACAAAG-----GAATTTTGGGCCCACTTTTGGATTCA 2034
Db 520 LysLysThrGluHisAlaAsnLeuLeuTyArgLysThrTyrcysIleSerGlnGlu 539
Qy 2035 GCAGTGTCTTCTCTGACTACTACTCTCCAGAAATTTCTTTTGAAGTACTACTGA 2094
Db 540 AlaGly---LeuProValIleIleLysGluThrIle-----LysIleSerAspAsp 555
Qy 2095 TTTACATGTGTGGATGCTCTACAAGCTCATGATCTACAGCTCGGAAGAATAAT--- 2151
Db 556 PheAspAlaLeuIleLysLeuSerIleProLysAspIleTyrcysIleSerHisGln 575
Qy 2152 -----CCTACTGTCTCTTATATATGTTGGTCTCTCAGTGCAGTGGTGAATATCGG 2205
Db 576 AlaIleProLeuIleValHisValTyrcysIleGlyProAsnAspGln-----AsnThr 592
Qy 2206 TTTAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGGTTGTA 2265
Db 593 LysGluAlaThrGlnIleGlyIleGluValValAlaSerAlaSerGlnAlaIle 612
Qy 2266 GTG---ATAGACACAGGGATCTGTCCAGGAGCTTAATTTGAAGCGCTTAA 2322
Db 613 LeuArgIleAspGlyArgGlySerGlyIleArgGlyTrpLysTyrcysIleAlaIleTy 632
Qy 2323 TATAAATGGGTCAATAGAAATGACGATCAGTGGAGGATCCCAATATCTAGCTTCT 2382
Db 633 GlyGlnLeuGlyThrValGluValGluAspGlnIleLysAlaIleValValLeuArg 652
Qy 2383 CGATAT---GATTTCAATGACTTAGATCGTGTGGGCATCCAGCGCTGGTCTATGAGA 2439
Db 653 LeuTyArgHisLeuLeuAspAlaArgValAlaValPheGlyTrpSerTyrcysIle 672
Qy 2440 TACCTCTCCGTGATGATTAATGCAGAGTTCAGAT---ATCTTCAGGGTGGTATTGCT 2496
Db 673 PheMetThrLeuSerMetValAsnGlnAlaProGlnPhePheLysCysAlaValSer 692
Qy 2497 GGGGCCCCAGTCTCTGTGATCTCTATGATACAGATACAGGAAGCTTATATGGGT 2556
Db 693 ValAlaProValThrAsnPheAlaTyrcysIleValAlaValValValValValVal 712
Qy 2557 CACCCGTGACCAAGTGAACGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAG 2616
```

```
Db 713 -----AspAlaProLeuGluSerTy-----SerAspValThrLysLeuAspAsn 728
Qy 2617 TTCCCTCTGACAAATCGTTTACTGTCTTACATGTTTCTCGATGAGAAATGTCAT 2676
Db 729 PheLysSer-----ThrArgLeuLeuMetHisGlyLeuLeuAspAsnValHis 746
Qy 2677 TTTGCCACATACAGTATATATCTAGTGTTTTGTAGTGGGCTGGAAAGCCATATGATTA 2736
Db 747 PheGlnAsnSerAlaIleLeuIleAspGluLeuGlnAsnArgGlyValAspPheAspLeu 766
Qy 2737 CAGATCTATCTCAGGAGACACAGCATAAGAGTTCCTGAATCGGAGAGATATATGAA 2796
Db 767 MetValTyProAsnGlnAlaHisSerLeuSerSerArgThrSer-----HisValVal 784
Qy 2797 CTGCATCTTTTTCACCTACCTCAAGAA 2823
Db 785 GlyLysMetThrHisPheLeuArgGln 793
RESULT 12
T25173
hypothetical protein T23F1.7a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25173
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25173
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-779 <WIL>
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN:CAB03411.1
A:Experimental source: clone T23F1
C:Genetics:
A:Gene: CESP:T23F1.7a
A:Map position: 5
A:Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1
C:Superfamily: dipeptidyl-peptidase IV
Alignment Scores:
Pred. No.: 3,39e-24 Length: 779
Score: 445.50 Matches: 193
Percent Similarity: 38.0% Conservative: 126
Best Local Similarity: 23.0% Mismatches: 292
Query Match: 8.0% Indels: 229
DB: 2 Gaps: 38
US-10-825-632-2 (1-3120) x T25173 (1-779)
Qy 445 CCTCATTCAGACAGATCTATTACCTTGCCATGCTCT-----GGTGAGAAC 489
Db 116 ProSerAlaAspArgLysTyrcysIlePheAlaMetMetAspHisAlaProAsnProGlyMetAsn 135
Qy 490 AGAGAAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGACGACGATCTTA 549
Db 136 ProGlnAsnGluThrPheHisLeuLysIleValAsnAsnGluArg----- 151
Qy 550 ATCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGCGATATGGAATGTAT 609
Db 152 -----IleThrTyArgPheGlyLeu--- 158
Qy 610 TCTCGAAGAAAGAACTATTAAAGAGAAAGAAACCGCATTTGGAACAGTCGGAATGCTTCT 669
Db 159 ---ArgLysGluLysVal-----IleGlnAla 167
Qy 670 TACGATTATCACAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAC 729
Db 168 PheLysTrpAsnGlyLysPheAsnAspPheValPheValGluSerAsnLysIleTyrcys 187
Qy 730 GTAAAAGATGGAGGGCCACAGGATTTAGGCACACCTTTAAGGCCCAATCTAGTGAA 789
Db 187 ----- 187
```

790	ACTAGTGTGCCAACATACCGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATT	849	Db	438	SerLeuPheSerTyrSer	-----LeuAlaAspGluSer	448	
188	GlnSerSerPro	-----	Qy	1876	TACTCAGATTTCTGCTGCATCAGT	-----CAGCAGTGTGACTCTTTTATAAGTAAG	1926	
850	GCTTTTATACATAGCAACGATATTGATGATATCTAACATCGTAACACAGAGAAGAAGAGA	909	Db	449	ArgAsnSerAlaTyrCysIleSerCysSerIleLeuAsnCysThrTrpAlaGlnAlaGln	-----	468	
192	-----GluGluGluGly	195	Qy	1927	TATAGTAAACAGAGAATCCACACTGTGTGCTCTCTTTACAGGATCTACAGTCTCTGAAGAT	1986		
910	CTCACTTATGTGCACAAATAGCTAGCAGCAACATCGAAGAGAGATCCAGATCAGCTGGAGTC	969	Db	469	MetAspAspGlnMetLysThrAlaIleValSerCysGlyGlyProAlaAlaProHisThr	488		
196	LeuThrArgValSerAsnGlyGlyGluHisThrValAsp	-----GlyLeu	Qy	1987	-----GACCCCAACTTGCACAAACAAAG	-----	2007	
970	GCTACCTTTGTTCTTCCAAAGAGAA	-----TTTGATAGATATTTCTGGCTATTGGTGGTCTCCA	1026	Db	489	AlaIleValAsnLeuThrArgMetAspSerAspLysLysThrGluHisAlaAsnLeuLeu	508	
211	PheAspTrpIleTyrGluGluGluIlePheGlyArgLysAspAlaMetTrpTrpSerThr	230	Qy	2008	-----GAATTTTGGGCCACCATTTTGGATTTCAGCAGGCTCTCTCTCTGACTATATCTCT	2061		
1027	AAAGCTGAACAACTCCACAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGAAAATGAT	1086	Db	509	TyrAspLysThrTyrGlnAsnArgValGluGluAlaGly	-----LeuProValIleIleLys	527	
231	Lys	-----GlyAspGlnLeuAlaTyrAlaSerTyrAspAsnHisLeu	244	Qy	2062	CCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGTATGGGATGCTCTACAAG	2121	
1087	GAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTCTTGTGGAAACAGAGGGCAGAT	1146	Db	528	GluThrIle	-----LysIleSerAspAspPheAspAlaLeuIleLysLeuSerIle	544	
245	ThrLysAsnValSerLeu	-----LysThrTyrHisArgLeuGluProTyrProIleAsp	262	Qy	2122	CCTCATGATCTACAGCTCGGAAAGAAATAT	-----CCTACTGTGCTGTTTCATATAT	2172
1147	---TCATTCCGTTATCTTAAACA	-----GGTACAGCAAAATCCTTAAAGTC	1188	Db	545	ProLysAspIleTyrAsnArgAspLysHisGlnAlaIleProLeuIleValHisValTyr	564	
263	ThrAsnPheHisTyrProLysThrPheAlaLysValLeuProThrTyrThrLeuSerIle	282	Qy	2173	GGTGTCTCTCAGGTCCAGTCTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTG	2232		
1189	ACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATCATAGATGTCATAGAT	1248	Db	565	GlyGlyProAsnAspGln	-----AsnThrLysGluAlaThrGlnIleGlyIleGlu	581	
283	TrpAsnLysLysThrGlu	-----GlnSerArgGlnLeuAspValGlnLeu	297	Qy	2233	ATAACCTAGCTCTCTAGTTATGCTGTTGTTAGTG	-----ATAGACAAACAGGGGATCTCTGT	2289
1249	AGGAACATAATCAACCTTTGAGATTTCTA	-----TTTGAAGGAGTT	1290	Db	582	GluValValAlaSerAlaSerGlnAlaIleLeuArgIleAspGlyArgGlySerGly	601	
298	LysAspSerLeuSerTyrHisTyrLeuLeuAlaValLysTrpLeuGluIleAsnGlyThr	317	Qy	2290	CACCGAGGCTTAAATTTGAGGCCCTTAAATATAAATGAGTCAAAATAGAAATGAC	2349		
1291	GAATATTTGCCAGAGCTGGATGACTCTGCTGAGGAAATATGCTTGGTCCATCTACTA	1350	Db	602	GlyArgGlyTrpLysTyrArgSerAlaIleTyrGlyGlnLeuGlyThrValGluValGlu	621		
318	GluGlnLeuValSerVal	-----TrpThr	334	Qy	2350	GATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATAT	-----GATTCATTCAGCTTAGAT	2406
1351	-----GATCGTCCAGACTCGCTACAGATAGTGTGCTCACCTGAA	1395	Db	622	AspGlnIleLysAlaIleLysValLeuLeuArgLeuTyrArgHisLeuLeuAspAlaArg	641		
335	ThrIleCysAspTrpAspThrAlaIleCysArgLeuGluPheGlu	-----	349	Qy	2407	CGTGTGGGCATCCACGGCTGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCAG	2466	
1396	TTATTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATGTCAGTGCCT	1455	Db	642	ArgValAlaValPheGlyTrpSerTyrGlyGlyPheMetThrLeuSerMetValAsnGlu	661		
349	-----	349	Qy	2467	AGGTCCAGAT	-----ATCTTCAGGGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTC	2523	
1456	GATTCTGTGACGCCCAATATCTATGAAGAAAACAACAGACATCTGGATAAATATCCAT	1515	Db	662	AlaProGluGlnPhePheLysCysAlaValSerValAlaProValThrAsnPheAlaTyr	681		
350	-----TyrLysTyrAlaSerLysArgTrpValThr	360	Qy	2524	TATGATACAGGATACACGGAACGTTTATATGGGTCCACCTGACCCAGAGAATGAACAGGGCTAT	2583		
1516	GACATCTTTCATGTTTTTCCCAAAAGTCACGAAGAGAAATTTGAGTTTATTTTCCCTCT	1575	Db	682	TyrAspAlaThrTyrThrGluArgTyrMetGly	-----AspAlaProLeuGluSerTyr	699	
361	AspAspPheHisSerIle	-----ThrSerPheGluAspThrLeuPhePheLeuLeuPro	378	Qy	2584	TACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTG	2643	
1576	GAATGCAAAACAGGTTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAAA	1635	Db	700	-----SerAspValThrLysLysLeuAspAsnPheLysSer	-----ThrArgLeuLeu	715	
379	-----HisAspLysArgAspAsnAlaPheGlnGlnValAlaSer	391	Qy	2644	CTCTTACATGTTTCTGGATGAGAATGTCCATTTTGCATACACATACCATGATATTTACTGAGT	2703		
1636	TATAACGATCCAGTGGTGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAGAG	1695	Db	716	LeuMetHisGlyLeuLeuAspAspAsnValHisPheGlnAsnSerAlaIleLeuIleAsp	735		
392	LeuArgLeuSerHisGlyGlnLeuArgThrProLys	403	Qy	2704	TTTTTAGTGAGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCCTCAGAGAGACACAGC	2763		
1696	GAGATAGCAATTACCAAGTGGTGAATGGAAATTTCTTGGCCGCGATGGATCTAATATCCAA	1755	Db	736	GluLeuGlnAsnArgGlyValAspPheAspLeuMetValTyrProAsnGlnAlaHisSer	755		
404	-----PheLeuAsnLeuGlyGluTyrAspValThrSerIleAsnGlyIleAsnLysGlu	421	Qy	2764	ATAAGAGTTCCTGAATCGGAGAACATATGAATCATCTCTTTTGCATCTACCTTCAAGAA	2823		
1756	GTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCCACCAAGACTCCCTTTTAGAGCAT	1815	Db	756	LeuSerSerArgThrSer	-----HisValValGlyLysMetThrHisPheLeuArgGln	773	
422	Thr	-----ArgThrIlePhePheHisAlaAlaProLysProSerHisArg	437	RESULT 13				
				IS4331				

dipeptidyl aminopeptidase like protein - human

C:Species: Homo sapiens (man)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C:Accession: I54331

R:Yokotani, N.; Doi, K.; Wentholt, R.J.; Wada, K.

Hum. Mol. Genet. 2, 1037-1039, 1993

A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related

A:Reference number: I54331; MUID:93372805; PMID:8103397

A:Accession: I54331

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-865 <RES>

A:Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:g306705; PIDN:

C:Superfamily: dipeptidyl-peptidase IV

Alignment Scores:

Pred. No.:	4.45e-24	Length:	865
Score:	444.00	Matches:	222
Percent Similarity:	37.0%	Conservative:	145
Best Local Similarity:	22.4%	Mismatches:	385
Query Match:	8.0%	Indels:	239
DB:	2	Gaps:	40

US-10-825-632-2 (1-3120) x I54331 (1-865)

QY	82	CGGCGCGCGCGCGAGGAAGCCACTGC-----AACGAGGACCGGAGTGGAGGCGG	132
DB	19	ProAlaProPro-GluAlaSerHisLeuLeuGlyGlyGlnGlyProGluGluAAspGlyG1	38
QY	133	CGGAGCATGAAGCGCGCGCGCGTCCATAGCGCA-----CGTCGGGA	177
DB	38	yAlaGlyAlaLysProLeuGlyProArgAlaGlnAlaAlaProArgGluAAspGlyG1	58
QY	178	CGTCCGCGCGCGCGCGGGAAGGAATGCAACATGG-----	217
DB	58	yGlyGlyGlyAlaGlyArgProArgPheGlnTyrGlnGlyArgSerAspGlyAs	78
QY	218	-----CAGCAGCAATCGGAACAGCAACAGCAGCTG	243
DB	78	pGluGluAAspGluLeuValGlySerAsnProProGlnArgAsnTrpLys-----	94
QY	244	GGTGTGAGATATTGGAAGCTGCGGACTGTGAGGAGAAATATTGAATCAGAGTCCGCT	303
DB	95	-----GlyI1	96
QY	304	AAATTGGAGCCTTTTATGTTGAGCGGTATCTCGAGTCAGCTTAAAGAGCTGCTG--	361
DB	96	eAlaIleAlaLeuValIleLeuValIleCysSerLeuIleValThrSerValIleLe	116
QY	362	-----CCGATACCGAAATATCATGCTACATGCTAAGGCTAAGGCCACCATGATTC	414
DB	116	uLeuThrProAlaGluAAsn-----	123
QY	415	ATGTTGTGAAGGAATGATCCAGATGACCTCATTCAGACAGAACTTATTCCTTGCC	474
DB	124	-----SerSerGlnLysLysLysValThrValG1	134
QY	475	ATGTCGTGTGAGACAGAG---AAATACACTGTTTATCTGAAATCCCAAACTATC	531
DB	134	uAspLeuPheSerGluAAspPheLysIleHis-----As	145
QY	532	AATAGACAGCAGCTTAATGCTCTCTGGAAGCCTCTTTGGATCTTTTTCAGGCAACA	591
DB	145	pProGlu-AlaLysTrpIleSerAspThrGluPheIleTyrArgGluGlnLysGlyThrV	165
QY	592	CTGACTATGCAATGTTATCTCGAAGGAAGAACTATTAAAG---GAAAGAAACGCATT	648
DB	165	alArgLeuTrpAsnValGluThrAsnThrSerThrVal-LeuIleGluGlyLysLysIle	184
QY	649	GGAACAGTCGGAATGCTTCTTACGATTATCACCAGGAAGTGGAACTTCTGTTT---	705
DB	185	GluSerLeuArgAlaIleArgTyrGluIleSerProAspArgGluTyrAlaLeuPheSer	204

QY	706	-----CAAGCGGTAGTGGAAATTTTATCAGTAAAGATGGAGGG	744
DB	205	TyrAsnValGluProIleTyrGlnHisSerTyrThrGlyTyrValLeuSerLysIle	224
QY	745	CCACAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAAC	804
DB	225	ProHisGlyAAspProGlnSerLeuAAspProGluValSerAsnAlaLysLeuGlnTyr	244
QY	805	ATACGGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGC	864
DB	245	AlaGlyTrpGlyProLys-----GlyGlnGlnLeuIlePheIlePheGlu	259
QY	865	AACGATATTTGGATATCTAACATCGTAACGAGGAAGAGGAGACTCACTTATGTGCAC	924
DB	260	AsnAsnIleTyrTyrCysAlaHisValGlyLysGlnAlaIleArgVal-----	275
QY	925	AATGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTCTC	984
DB	276	-----ValSerThrGlyLysGluGlyValIleTyrAsnGlyLeuSerAspTrpLeuTyr	293
QY	985	CAAGAAGAA---TTTGATAGATATCTCGCTATTGCTGTGTCCTCAAAAGCTGAAACAACT	1041
DB	294	GluGluGluLeuLeuLysThrHisIleAlaHisTrpTrpSerProAspGly-----	310
QY	1042	CCAGTGGTGTAAATTTTGAATTTCTATATGAAGAAATATGAATCTGAGTGGAA	1101
DB	311	-----ThrArgLeuAlaTyrAlaAlaIleAsnAspSerArgValPro	324
QY	1102	ATTATTCATGTT-----ACATCCCTATGTTGGAAACAGGAGGCGAGATCATTC	1152
DB	325	IleMetGluLeuProThrTyrThrGlySerIleTyrProThr-----ValLysProTyr	342
QY	1153	CGTTATCCTAAACAGGTACAGCAATCCTAAGTCACTTTTAAGATGTCAGAAATATG	1212
DB	343	HisTyrProLysAlaGlySerGluAsnProSerIleSerLeuHis-----Val	358
QY	1213	ATTGATCCTGAAGAGGATCATAGATGTCATAGTAAGAACTAATTAACCTTTTCAG	1272
DB	359	IleGlyLeuAsnGlyProThrHisAspLeu-----GluMetProProAspAsp	375
QY	1273	ATTCTATTGAAGAGTTGAATATATATGTCAGAGCTGGAGTGCCTCCTGAGGAAATAT	1332
DB	376	ProArgMetArgGluTyr---TyrIleThrMetValLysTrpAlaThrSerThrLysVal	394
QY	1333	GCT-----TGTTCATCTCTAGTACCTCCAGACTCCGCTACAGATAGTGTGATC	1386
DB	395	AlaValThrTrp-----LeuAsnArgAlaGlnAsn-----Val	405
QY	1387	TCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAAGCAGACAGACTCAT	1446
DB	406	Ser-----IleLeu	408
QY	1447	TCAGTGCCTGATTTCTGTGAGCCACCAATATTC-----TATGAAAGAAACACAGACATC	1500
DB	409	ThrLeuCysAspAlaThrThrGlyValCysThrLysHisGluAspGluSerGluAla	428
QY	1501	TGGATAATATCCATGACATCTTTCATGTTTTCCTCCCAAGTCCAGAGAGAGAAAT---	1557
DB	429	TrpLeuHis-----ArgGlnAsnGluProValPhe	439
QY	1558	-----GAGTTTATTTTGGCTCTGAATGCAAAACAGAGTTTCCGT---CAT	1599
DB	440	SerLysAspGlyArgLysPhePheIleArgAlaIleProGlnGlyGlyArgGlyLys	459
QY	1600	TTATACAAAATTACATCTATTTTAAAGGAAGAAACAAATATAAACCATCCAGTGGGGCTG	1659
DB	460	PheTyrHisIleThr-----ValSerSerSerGlnProAsnSerSerAsnAspAsnIle	477
QY	1660	CCTGCTCCAAGTGAATTTCAAGTGTCTATCAAAAGAGGAGATAGCAATTCACAGTGGTGA	1719
DB	478	Gln-----SerIleThrSerGlyAsp	484
QY	1720	TGGAGATTTCTGCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTA	1779


```
Db 307 GluMetThrProAspAspProArgMetArgGluTyr---TyrIleThrMetValLys 325
|||::: ||| ::: |||||
1312 TGACCTCCTGAGGGAATATGCT-----TGGTCCATCTCTAGATCCCTCCAGACT 1365
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
326 TrpAlaThrSerThrLysValAlaValAsnTrp-----LeuSerArgAlaGlnAsn 342
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1366 CGCTACAGATAGTGTGATCTCACCCTGAATATTATATCCAGTAGAAGATGATGTATG 1425
::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
343 -----ValSer----- 344
1426 GAAGGCAGAGACTCATTTGAGTCAGTCGCTGATTTCTGTGCGCCACTAATTATC----- 1479
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
345 -----IleLeuThrLeuCysAspAlaThrThrGlyValCysThrLysLys 359
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1480 TATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTCCTCCCAA 1539
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
360 HisGluAspGluSerGluAlaTrpLys----- 368
1540 AGTCACGAAGAGAAATTGAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTCCCTCAT 1599
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
369 ---HisArgGlnAsnGluProValPheSer-----LysAspGly---ArgLys 383
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1600 TTATCAAAATTTACATCTATTTTAAAG-----GAAAGCAAAATATAACCATCCAGTGT 1653
::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
384 PhePheValArgAlaIleProGlnGlyGlyPheThrHisIleThrVal 403
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1654 GGCTCCCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATACCAGT 1713
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
404 SerSerSerGlnProAsnSerSerAsnAspAlaIleGln-----SerIleThrSer 420
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1714 GGTGAATGGAAAGTCTTTGGCCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG 1773
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
421 GlyAspTrpAspValThr-----LysIleLeuSerTyrAspGluLysArgSer 436
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
437 GlnIleTyrPheLeuSerThrGluAspLeuProArgArgGlnLeuTyrSerAlaSer 456
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1834 TAGCTAAATCCTGGAGGTGACAGG-----CTGACTGACCTGGC 1875
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
457 ThrVal-----GlySerPheAsnArgGlnCysLeuSerCysAspLeuValAspCys 474
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1876 TACTCACATCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTATAAGTAAAGTATAGTAAC 1935
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
475 ThrTyrPheSerAlaSerPheSerProGlyAlaAspPhePheLeuLeuLysCysGluGly 494
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1936 CAGAAGATCCACACTGTGTGTCCTTTTACAAGTCATCAAGTCCTGAAGATGACCCCACT 1995
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
495 ProGlyValPro---ThrValSerValHisAsnThrThrAspLysLysLysMetPheAsp 513
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1996 TGCMAAACAAAGGA-----TTTTGGCCACCATTTGGATTTCAGCAGTCTCTTCCTGAC 2052
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
514 LeuGluThrAsnGluHisValGlnLysAlaIleSerAspArgGlnMetProLysValGlu 533
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2053 TATACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATG 2112
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
534 TyrArgLysIleGlu-----ThrAspAspTyrAsnLeuProIleGln 547
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2113 CTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATAT 2172
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
548 IleLeuLysProAlaThrPheThrAspThrAlaHisTyrProLeuLeuLeuValValAsp 567
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2173 GGTGTCTCAGTCGTCAGTGTGTGAATATCGGTTTAAAGGATCAAGTATTTCCGCTTG 2232
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
568 GlyThrProGlySerGlnSerValAlaGluLysPhe-----AlaValThrTrpGluThr 585
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2233 AATACCTAGCTCTCTAGTGTATGCTGTGTAGTGATAGACACAGGGATCTGTCAC 2292
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
586 valMetValSerSerHisGlyAlaValValLysCysAspGlyArgGlySerGlyPhe 605
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2293 CGAGGCTTTAAATTTGAAGCGCCTTTAAATATATAAATGGGTCAAAATAGAAATTCAGCAT 2352
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
```

```
Db 606 GlnGlyThrArgLeuLeuHisGluValArgArgLeuGlySerLeuGluGluLysAsp 625
2353 CAGGTGGAAGGACTCCAATATATCTCTCGATATGATTTTATTGATCTAGATCGTGTG 2412
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
626 GlnMetGluAlaValArgValMetLeu---LysGluProTyrIleAspLysThrArgVal 644
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2413 GGCATCACGGCTGCTCTATCGAGATACCTCTCCCTGATGGCATTAATCCAGAGTCA 2472
::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
645 AlaValPheGlyLysAspTyrGlyGlyTyrLeuSerThrTyrLeuLeuProAlaLysGly 664
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2473 GAT-----ATCTTCAGGGTGTCTATTGCTGGGCCCAGTCACCTCTCTGGATC 2520
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
665 AspGlyGlnAlaProValPheSerCysGlySerAlaLeuSerProIleThrAspPheLys 684
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2521 TTCTATGATACAGATACACGGAAGCTTATATGGGTCCCTCCCTCAGCAAGTGAACAGGCG 2580
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
685 LeuTyrAlaSerAlaPheSerGluArgTyrLeuGlyLeuHisGlyLeuAspAsnArgAla 704
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAAGTTCCTCTGAACCAAAATCGTTTA 2640
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
705 TyrGluMetAlaLysValAlaHisArgValSerAlaLeu-----GluGlyGlnGlnPhe 722
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2641 CTGCTCTTACATGTTTCTCTGATGAGAATGCCATTTTGCACATACCAGTATATTACTG 2700
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
723 LeuValIleHisAlaThrAlaAspGluLysIleHisPheGlnHisThrAlaGluLeuIle 742
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2701 AGTTTTTATGTCAGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGACAC 2760
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
743 ThrGlnLeuIleLysGlyLysAlaAsnTyrSerLeuGlnIleTyrProAspGluSerHis 762
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2761 AGCATAAGAGTTCCTGAATCGGGAGAACATTATGACATGTCATCTTTTGCACACTCTCAA 2820
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
763 TyrPheSerSerAlaAlaLeuGlnGlnHisLeuHisArgSerIleLeuGlyPhePheVal 782
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
QY 2821 GAA 2823
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db 783 Glu 783
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
RESULT 15
168600
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: I68600
R:Yokotani, N.; Doi, K.; Wenthhold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I68600
A:Molecule type: mRNA
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-803 <RES>
A:Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:g306707; PIDN:AAA35761.1; PID
C:Superfamily: dipeptidyl-peptidase IV
Alignment Scores:
Pred. No.: 2,99e-23 Length: 803
Score: 432.50 Matches: 185
Percent Similarity: 39.4% Conservative: 127
Best Local Similarity: 23.4% Mismatches: 326
Query Match: 7.8% Indels: 154
DB: 2 Gaps: 32
US-10-825-632-2 (1-3120) x I68600 (1-803)
QY 585 GGCAACACCTGAGCTATGGAATGTATTCTCGAGAGAACACTATTAGA---GAAAGAAA 641
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db 101 GlyThrValArgLeuTrpAsnValGluThrAsnThrSerThrVal-LeuIleGluGlyLy 120
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
642 ACGCATTGGACAGTCGGAATTGCTTCTTACGATTATCACCAGAGTGGAACTTCT 701
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db 120 slyIleGluSerLeuArgAlaIleArgTyrGluIleSerProAspArgGluTyrAlaLe 140
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
```


Db 752 rSerLeuGlnIleTyrProAspGluSerHisTyrPheThrSerSerSerLeuLysGlnHi 772
QY 2790 TTATGAAGCTGCACTCTTTTGCACTACCTTCAAGAA 2823
Db 772 sLeuTyrArgSerIleIleAsnPhePheValGlu 783

Search completed: April 14, 2006, 11:34:39
Job time : 195.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 14, 2006, 10:51:36 ; Search time 137.6 Seconds
(without alignments)
3199.489 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5552
Sequence: 1 aagtgtaaagctccgagg.....agaattactataaaaaaaa 3120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xl-p
-O=/abs/ABSSWEB.spool/US10825632/runat_14042006_092349_3432/app_query.fasta_1
-DB=UniProt -QMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10825632 @CGN 1 1 466 @runat_14042006_092349_3432 -NCP0=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4726	85.1	898	1 DPP8 HUMAN	Q6vix1 homo sapien
2	4559.5	82.1	892	1 DPP8 MOUSE	Q80ya7 mus musculus
3	2870	51.7	863	1 DPP9 HUMAN	Q86ti2 homo sapien
4	2833	51.0	862	1 DPP9 MOUSE	Q8bvg4 mus musculus
5	2829.5	51.0	923	2 Q48M6 TETNG	Q48m6 tetraodon n
6	2707.5	48.8	847	2 Q6GR22 XENLA	Q6gr22 xenopus lae
7	1643.5	29.6	508	2 Q75273 HUMAN	Q75273 homo sapien
8	1641	29.6	886	2 Q7QBK1 ANOGA	Q7qbk1 anopheles g
9	1624.5	29.3	1053	2 Q9VC20 DROME	Q9vc20 drosophila
10	1607.5	29.0	1113	2 Q9VC19 DROME	Q9vc19 drosophila
11	1567	28.2	740	2 Q5TKK8 ANOGA	Q5tkk8 anopheles g
12	1321	23.8	621	2 Q7PTT8 ANOGA	Q7ptt8 anopheles g
13	1142.5	20.6	557	2 Q5TXJ2 ANOGA	Q5txj2 anopheles g
14	1124.5	20.3	803	2 Q54U01 DICDI	Q54u01 dictyosteli
15	1049	18.9	432	2 Q75868 HUMAN	Q75868 homo sapien
16	924.5	16.7	927	2 Q965K3 CAEEL	Q965k3 caenorhabdi

17	922.5	16.6	931	2	044987 CAEEL	044987 caenorhabdi
18	867.5	15.6	895	2	061CU7 CAEEL	061cu7 caenorhabdi
19	846	15.2	763	2	08EAB7 SHEON	08eab7 shewanella
20	842	15.2	738	2	09A6E0 CAUCR	09a6e0 caulobacter
21	837.5	15.1	746	2	09FNF6 ARATH	09fnf6 arabidopsis
22	792.5	14.3	745	2	06F3I7 PSED	06f3i7 pseudomonas
23	786	14.2	596	2	06K880 ORISA	06k880 oryza sativ
24	779	14.0	743	2	05QX36 IDILO	05qx36 idiomarina
25	759	13.7	745	2	05H5W8 XANOR	05h5w8 xanthomonas
26	754	13.6	741	2	057822 XANWA	057822 xanthomonas
27	751.5	13.5	751	2	04UPD3 XANCP	04upd3 xanthomonas
28	751.5	13.5	751	2	08P3V8 XANCP	08p3v8 xanthomonas
29	746	13.4	766	2	04TNE1 SPHEN	04tnel erythrobact
30	744.5	13.4	757	2	08PFD7 XANAC	08pfd7 xanthomonas
31	727	13.1	552	2	08GUJ7 ARATH	08guj7 arabidopsis
32	724.5	13.0	720	2	05NNM8 ZYMMO	05nnm8 zymomonas m
33	722.5	13.0	749	2	07NEK8 GIOVI	07nek8 gloeobacter
34	672	12.1	736	2	08A028 BACTN	08a028 bacteroides
35	670	12.1	739	2	05FQI6 GLUOX	05fqiy6 gluconobact
36	655.5	11.8	850	2	06H9E3 TRYP	06h9e3 trypanosoma
37	651	11.7	736	2	05LGU5 BACFN	05lgus bacteroides
38	648	11.7	736	2	064XP9 BACFR	064xp9 bacteroides
39	643	11.6	852	2	04QIH9 LEIMA	04qih9 leishmania
40	620	11.2	711	2	047900 FLAME	047900 flavobacter
41	606	10.9	809	2	04NVV1 GDELT	04nvv1 anaeromyxob
42	595.5	10.7	778	2	051ZF7 MAGGR	051zf7 magnaporthe
43	587	10.6	769	2	05QVY7 IDILO	05qvy7 idiomarina
44	582.5	10.5	771	2	042812 ASPOR	042812 aspergillus
45	578.5	10.4	731	2	06L872 PREIN	06l872 prevotella

ALIGNMENTS

RESULT 1

DPP8_HUMAN	STANDARD;	PRT; 898 AA.
ID	DPP8 HUMAN	Q724D3; Q724E1; Q81WC7; Q8NEM5; Q96JX1; Q9HBM2;
AC	Q6VIX1; Q724C9; Q724D3; Q724E1; Q81WC7; Q8NEM5; Q9HBM2;	
AC	Q9HBM3; Q9HBM4; Q9HBM5; Q9NKF4;	
DT	13-SEP-2005 (Rel. 48, Last sequence update)	
DT	13-SEP-2005 (Rel. 48, Last sequence update)	
DT	13-SEP-2005 (Rel. 48, Last sequence update)	
DE	Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DPP8)	
DE	(Prolyl) dipeptidase DPP8 (Dipeptidyl peptidase IV-related protein 1)	
DE	(DPP-1)	
GN	Name=DPP8; Synonyms=DPP1; ORFNames=MSTP097, MSTP135, MSTP141;	
OS	Homo sapiens (Human)	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;	
OC	Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION, CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION, AND SUBCELLULAR LOCATION.	
RC	TISSUE=Placenta;	
RE	MEDLINE=20467194; PubMed=11012666;	
RA	Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W., Gorrell M.D.;	
RA	"Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8.;"	
RL	Eur. J. Biochem. 267:6140-6150(2000).	
RN	[2]	
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME REGULATION, BIOPHYSICOCHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.	
RC	TISSUE=Testis.;"	
RE	PubMed=12662155; DOI=10.1042/BJ20021914;	
RA	Qi S.Y., Riviere P.J., Trojanar J., Junien J.-L., Akinsanya K.O.;	
RA	"Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases.;"	
RL	Biochem. J. 373:179-189(2003).	

[3] NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
 RP TISSUE=Testis;
 RA Shu J.H., Zhou Z.M., Li J.M.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 [4] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE
 RP SEQUENCE [LARGE SCALE MRNA] OF 211-898 (ISOFORM 2).
 RC TISSUE=Hepatoma, and Placenta;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa K.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotoh T.,
 RA Kusanu J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togaiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yohikawa Y., Matsunawa H., Ichihara T., Shichata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Misoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Koniya M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Fujimori Y., Koniya M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai F., Takahashi Y., Yamashita R.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Sugano S.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 [5] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
 RP TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lochell N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Hellon E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Merris M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [6] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 561-898.
 RP TISSUE=Aorta;
 RA Zhao B., Xu H.S., Tong Y.K., Sheng H., Qin B.M., Liu Y.Q., Liu B.,
 RA Wang X.Y., Zhang Q., Song L., Gao Y., Zhang C.L., Ye J., Ji X.J.,
 RA Liu B.H., Lu H., Chen J.Z., Cai M.Q., Zheng W.Y., Teng C.Y., Liu Q.,
 RA Yu L.T., Lin J., Gong Q., Zhang A.M., Gao R.L., Hui R.T.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 [7] MUTAGENESIS OF GLU-275; SER-755; ASP-833 AND HIS-865, CATALYTIC
 RP ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, AND SUBCELLULAR LOCATION.
 RX PubMed=12534281; DOI=10.1021/bi026946s;
 RA Ajami K., Abbott C.A., Obradovic M., Gysbers V., Kaehne T.,
 RA McCaughan G.W., Gorrell M.D.;
 RT "Structural requirements for catalysis, expression, and dimerization
 RT in the CD26/bpiv gene family.";
 RL Biochemistry 42:694-701(2003).
 [8] BIOPHYSICOCHEMICAL PROPERTIES.
 RP PubMed=15039077; DOI=10.1016/j.bj.2003.12.019;
 RA Chen Y.-S., Chien C.-H., Goparaju C.M., Hsu J.T.-A., Liang P.-H.,
 RA Chen X.;
 RT "Purification and characterization of human prolyl dipeptidase DPP8 in
 RT Sf9 insect cells.";
 RL Protein Expr. Purif. 35:142-146(2004).
 [9] ENZYME REGULATION.
 RP PubMed=15664838; DOI=10.1016/j.bmc.2004.11.023;
 RA Jiaang W.-T., Chen Y.-S., Hsu T., Wu S.-H., Chien C.-H., Chang C.-N.,
 RA Chang S.-P., Lee S.-J., Chen X.;
 RT "Novel isoindoline compounds for potent and selective inhibition of
 RT prolyl dipeptidase DPP8.";
 RL Bioorg. Med. Chem. Lett. 15:687-691(2005).
 CC -1- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
 CC dipeptides from proteins having a Pro or Ala residue at position
 CC 2. May play a role in T-cell activation and immune function. Yaa-|
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|
 CC zaa, from a polypeptide, preferentially when Yaa is Pro, provided
 CC zaa is neither Pro nor hydroxyproline.
 CC -1- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
 CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride
 CC (AEPF), and by di-isopropylfluorophosphate. Specifically inhibited
 CC by isoindoline derivatives.
 CC -1- BIOPHYSICOCHEMICAL PROPERTIES:
 CC Kinetic parameters:
 CC KM=208 uM for Ala-Pro-AMC;
 CC KM=130 uM for Ala-Pro-AFC;
 CC KM=120 uM for H-Ala-Pro-pNa;
 CC KM=1420 uM for H-Ala-Ala-pNa;
 CC KM=310 uM for H-Arg-Pro-pNa;
 CC KM=2050 uM for H-Asp-Pro-pNa;
 CC KM=480 uM for H-Gly-Pro-pNa,
 CC pH dependence:
 CC Optimum pH is 7.4-8.5. Little activity below pH 6.5;
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=6;
 CC Name=1;
 CC IsoId=Q6V1X1-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q6V1X1-2; Sequence=VSP_013864;
 CC Name=3;
 CC IsoId=Q6V1X1-3; Sequence=VSP_013860;
 CC Name=4;
 CC IsoId=Q6V1X1-4; Sequence=VSP_013860, VSP_013862;
 CC Name=5;
 CC IsoId=Q6V1X1-5; Sequence=VSP_013863;
 CC Name=6;
 CC IsoId=Q6V1X1-6; Sequence=VSP_013861;
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
 CC testis, placenta, prostate, muscle and brain.
 CC -1- INDUCTION: In activated T-cells.
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC -1- CAUTION: Ref.4 (BAA91059) sequence differs from that shown due to
 CC a frameshift in position 486.
 CC -1- CAUTION: Ref.6 (AAQ13650 and AAQ13623) sequences differ from that
 CC shown due to several frameshifts.

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use as long as its content is in no way modified and this statement is not removed.

CC
CC
CC
DR EMBL; AF221634; AAG29766.1; -; mRNA.
DR EMBL; AF221635; AAG29767.1; -; mRNA.
DR EMBL; AF221636; AAG29768.1; -; mRNA.
DR EMBL; AF221637; AAG29769.1; -; mRNA.
DR EMBL; AY172659; AA017261.1; -; mRNA.
DR EMBL; AY34202; AAQ3887.1; -; mRNA.
DR EMBL; AK000290; BAA91059.1; ALT_FRAME; mRNA.
DR EMBL; AK027826; BAB55395.1; ALT_INIT; mRNA.
DR EMBL; BC030688; AAH30688.3; -; mRNA.
DR EMBL; BC040203; AAH40203.1; ALT_INIT; mRNA.
DR EMBL; AF176779; AAQ13657.1; ALT_INIT; mRNA.
DR EMBL; AF175225; AAQ13650.1; ALT_FRAME; mRNA.
DR EMBL; AF173382; AAQ13623.1; ALT_FRAME; mRNA.

Alignment Scores:

Pred. No.: 0 Length: 898
Score: 4726.00 Matches: 886
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.1% Indels: 0
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x DPP8_HUMAN (1-898)

QY 202 GGAAATGCAATGGCAGCAGCAATGAAACAGACAGCTGGTGTGAGATATTGAA 261
DB 13 GlyLysCysAsnMetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGlu 32
QY 262 ACTGGCGACTGTGAGGAGAAATTTGAATCACAGATCGGCTAAATGGAGCCTTTTAT 321
DB 33 ThrAlaAspCysGluGlnuAlaMetGlnSerGlnAspArgProLysLeuGluProPheTyr 52
QY 322 GTTGACGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACCCAGAAATATCAT 381
DB 53 ValGluArgTyrSerTyrPheGlnLeuLysLeuLeuAlaAspThrArgLysTyrHis 72
QY 382 GGCTACATGATCGCTAAGGCCACCATGATTTCATGTTGTGAGAGGAATGATCAGAT 441
DB 73 GlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAsp 92
QY 442 GGACCTCATPTACAGCAATCTATTACCTGCCATGTCTGGTGAACAGAGAAATACA 501
DB 93 GlyProHisSerAspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThr 112
QY 502 CTGTTTATTCGAAATCCCAACTATCAATAGCAGCAGCTTAAATGCTCTCTGG 561
DB 113 LeuPheTyrSerGluLeuProLysThrIleAsnArgAlaValLeuMetLeuSerTyr 132
QY 562 AAGCCTCTTTTGGATCTTTTTCAGGCAACATCGGACTATGGAATGTATTCTCAGAGAA 621
DB 133 LysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGlu 152
QY 622 GAACATTAAGAGAAAGAAACGCAATTCGAAACAGTCGGAATGCTTTTACGATATATCAC 681
DB 153 GluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHis 172
QY 682 CAAGGAAGTGGAAACATTTCTGTTTCAAGCCGTAGTGAATTTATCAGTAAAGATGGA 741
DB 173 GlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGly 192
QY 742 GGGCCACAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCC 801
DB 193 GlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysPro 212
QY 802 AACATACGGATGGATCCAAATATTATGCCCGCTGATCCAGACTGGATGCTTTTATACAT 861
DB 213 AsnIleArgMetAspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHis 232
QY 862 AGCAACGATATTGATATCTAACTCGTTAACCGAGAGAAAGGAGACTCACTTATGTG 921

DB 233 SerAsnAspIleTyrPheSerAsnIleValThrArgGluGluArgArgLeuThrTyrVal 252
QY 922 CACAATGAGCTAGCCACCAATCGAAGAAGATCCAGATCAGCTGAGTCGCTACTTTGTT 981
DB 253 HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal 272
QY 982 CTCCTAAGAAAGAAATTTGATAGATATTCGGCTATTGGTGGTGTCCAAAAGCTGAAACAAC 1041
DB 273 LeuGlnGluLupheAspArgTyrSerGlyTyrTyrPheCysProLysAlaGluThrThr 292
QY 1042 CCAGTGGTGGTAAATTTCTTAGAATCTTATGAGAAATATGATGAATCTGAGGCTGAA 1101
DB 293 ProSerGlyGlyLysIleLeuArgIleLeuTyrGluGlnAsnAspGluSerGluValGlu 312
QY 1102 ATTATTATGTTTACATCCCTTATGTTGAAAACAAGAGGGCAGATTCATTCCTGTTATCCT 1161
DB 313 IleIleHisValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrPro 332
QY 1162 AAAACAGGTACAGCAAAATCCTTAAGTCACTTTTAAGATGTCAGAAATATGATTGATCT 1221
DB 333 LysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAla 352
QY 1222 GAAGGAAGCATCATAGATGTCATAGATAAGGAACCTTAATTCACCTTTTGAGATTTTAT 1281
DB 353 GluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPhe 372
QY 1282 GAAGGAGTTGAATATATATCCAGAGCTGAGTGGACTCTCGAGGAAATATGCTTGGTCC 1341
DB 373 GluGlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSer 392
QY 1342 ATCCTACTAGATGCTCCAGACTCGCTACAGATAGTAGTGTGATCTCATCTGAATTTAT 1401
DB 393 IleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPhe 412
QY 1402 ATCCAGTAGAAGATGATGTTATGGAAGGAGAGACCTCATTGAGTCAGTGCCTGATTCT 1461
DB 413 IleProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSer 432
QY 1462 GTGACGCCCATTAATATCTATGAAGAAACACAGACATCTGGATAAATATCATGACATC 1521
DB 433 ValThrProLeuIleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIle 452
QY 1522 TTTTCATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTCCTCTGCAATGC 1581
DB 453 PheHisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCys 472
QY 1582 AAAACAGGTTTCGTCATTTATACAAATATCATCTATTTTAAAGGAAGCAATATATAA 1641
DB 473 LysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLys 492
QY 1642 CGATCCAGTGGTGGCTGCTGCTCCCAAGTGAATTCAGTGTCTCATCAAGAGAGAGATA 1701
DB 493 ArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluLeu 512
QY 1702 GCAATTAACAGTGTGATGGAAGTTCCTTGGCCGCGCATGATCTTAATATCAAGTTCAT 1761
DB 513 AlaIleThrSerGlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAsp 532
QY 1762 GAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGATCACCTG 1821
DB 533 GluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu 552
QY 1822 TACGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTCAGTCGCGCTGCTACTCA 1881
DB 553 TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer 572
QY 1882 CATCTTGTGCTGATCAGTCAGCTGCTGACTCTCTTTTATAGTAAGTATAGTAAACAGAG 1941
DB 573 HisSerCysLysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLys 592
QY 1942 AATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGC 2001
DB 593 AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 612

```

2002 ACAAGGAATTTTGGGCCACCATTTTGGATTACAGAGGTCCTCTCTGACTATATCTCT 2061
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
613 ThrLysGluPheThrPalaThrIleGluAspSerAlaGlyProLeuProAspGlyThrPro 632
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2062 CCAGAAATTTCTCTTTTGAAGTACTCTGGATTTACATGTATGGGATGCTCTACAG 2121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
633 ProGluIlePheSerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLys 652
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2122 CTTCTATGATCTACAGCCTGGAAAGAAATATCTCTCTGTCTGTCTATATATGTTGCTCT 2181
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
653 ProHisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyPro 672
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2182 CAGGTGCGATTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTA 2241
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
673 GlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeu 692
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2242 GCCTCTCTAGGTATGTGGTTGTAGTGTATGATAGACACAGGGGATCTCTGTCACCGAGGGCTT 2301
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
693 AlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeu 712
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2302 AAATTTGAAGGCGCTTTAAATATAAATGGTCAATAGAAATTCAGCATCAGGTGGAA 2361
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
713 LysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspGlnValGlu 732
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2362 GGACTCCATATCTAGCTTCCTCGATATGATTTCAATTCAGCTAGATCGTGGGCATCCAC 2421
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
733 GlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHis 752
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2422 GCCTGCTCTATGGAGGATCTCTCTCTGATGGCATTATGACAGAGGTTCAGATATCTTC 2481
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
753 GlyTyrPserTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePhe 772
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2482 AGGGTGTCTATTGCTGGGGCCCGCCAGTCTCTGTGGATCTTCTATGATACAGATACAG 2541
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
773 ArgValAlaIleAlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThr 792
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2542 GAACGTTATATGGGTCCACCTGACCAAGATGACAGAGGTCTATTCTAGGATCTGTGGCC 2601
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
793 GluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAla 812
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2602 ATGCAAGCAGAAAGTCTCCCTCTGACCAACCAATCGTTTACTGCTCTTACATGTTTCTCTG 2661
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
813 MetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeu 832
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2662 GATGAGATGTCCATTTTGCACATACACAGTATATTACTGAGTTTTTTAGTGAGGCTGGA 2721
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
833 AspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGly 852
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2722 AAGCCATATGATTTACAGATCTATCCTCAGGAGACACAGCATAAGAGTTCCTGAATCG 2781
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
853 LysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSer 872
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2782 CGAGACATTTATGACTGCATCTTTTGCACCTTACCTTCAAGAAACCTTGGATCAGCTATT 2841
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
873 GlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIle 892
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2842 GCTGCTCTAAAGTGTATA 2859
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
893 AlaAlaLeuLeuValIle 898
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
DPP8_MOUSE
ID DPP8_MOUSE
AC Q80YA7; Q9D4G6; PRT; 892 AA.
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DE 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII)
DE (DPP8).
GN Name=Dpp8;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=27534683; PubMed=12466851; DOI=10.1038/nature01266;
RX Okazaki Y., Furuno M., Kasukawa T., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Nikaido I., Osato N., Saito K., Tazaki K., Nishikawa S., Gojobori T.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kradnoff A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Brad D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszynski B., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Sasaki D., Shibata K., Shinagawa A.,
RA Birney E., Hayashizaki Y., Waterston R., Lander E.S., Rogers J.,
RA
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [2]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Brain;
RC MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
PROSPHYLATION SITE TYR-325
RX PubMed=14723942; DOI=10.1074/mcp.D300003-MCP200;
RA Shu H., Chen S., Bi Q., Mumbly M., Brekken D.L.,
RT "Identification of phosphoproteins and their phosphorylation sites in
RT the WEHI-231 B lymphoma cell line."
RL Mol. Cell. Proteomics 3:279-286 (2004).
CC -1- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a pro or Ala residue at position
CC 2. May play a role in T-cell activation and immune function (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
```


NP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 272-863 (ISOFORM 2), AND
 PP PARTIAL NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 5).
 RC TISSUE=Glial tumor, Ovary, Spleen, and Trachea;
 CC PubMed=14704039; DOI=10.1038/ngl285;
 CC Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 CC Wakamatsu A., Hayashi K., Sato H., Negai K., Kimura K., Makita H.,
 CC Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 CC Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 CC Negahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 CC Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 CC Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 CC Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 CC Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 CC Fujimori K., Tanai H., Kimata S., Watanabe M., Hirakata S., Chiba Y.,
 CC Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 CC Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 CC Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
 CC Imose N., Muesashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 CC Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 CC Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 CC Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 CC Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 CC Yamazaki Y., Komiya M., Tashiro H., Tanigami A., Fujiwara T.,
 CC Fujimori Y., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 CC Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 CC Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Serba T.,
 CC Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 CC Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 CC Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 CC Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 CC Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 CC "Complete sequencing and characterization of 21,243 full-length human
 CC cDNAs."; [6]
 CC NT Genet. 36:40-45(2004).
 CC [16]
 CC NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 209-863 (ISOFORM 4), AND
 CC NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 298-863 (ISOFORM 2).
 CC TISSUE=Melanoma;
 CC The German cDNA consortium;
 CC Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
 CC dipeptides from proteins having a Pro or Ala residue at position
 CC 2.
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-[-
 CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
 CC Zaa is neither pro nor hydroxyproline.
 CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
 CC (2-aminoethyl)benzenesulphonyl fluoride (AEDSF), and by di-
 CC isopropyl fluorophosphate.
 CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
 CC Kinetic parameters:
 CC KM=161 uM for Ala-Pro-AMC;
 CC KM=180 uM for Ala-Pro-AFC;
 CC pH dependence:
 CC Optimum pH is 7.5-8.5. Little activity below pH 6.5;
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; cytosol.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=5;
 CC Name=1; Synonyms=Short;
 CC IsoId=Q86T12-1; Sequence=Displayed;
 CC Name=2; Synonyms=Long;
 CC IsoId=Q86T12-2; Sequence=VSP_013865;
 CC Note=Incomplete sequence;
 CC Name=3;
 CC IsoId=Q86T12-3; Sequence=VSP_013867, VSP_013868;
 CC Name=4;
 CC IsoId=Q86T12-4; Sequence=VSP_013869;
 CC Name=5;
 CC IsoId=Q86T12-5; Sequence=VSP_013865, VSP_013866;
 CC Note=Incomplete sequence;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
 CC liver, heart and muscle, and lowest levels in brain

[illegible]

QY 436 CCAGATGACCTATTACAGACAGAACTATTACCTGCTGCCATGCTGGTGAACAGAGAA 495
Db 64 GluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyrGlySerArgGlu 83
QY 496 AATACACGTGTTTATTCGAATTCACCAAACTATCAATACAGACAGCTCTTAATGCTC 555
Db 84 AnSerLeuLeuSerGlyLeuProLysValArgLysGluAlaLeuLeuLeuLeu 103
QY 556 TCTTGAAGCCTCTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGATTCGA 615
Db 104 SerTyrLysGlnMetLeuAspHisPheGlnAlaThrProHisHisGlyValTyrSerArg 123
QY 616 GAAGAAGAACTATTAAGAGAAAGAAACCGATGGACAGCTCGGAATCTCTTACCAT 675
Db 124 GluGluGluLeuLeuArgGluArgLeuGlyValPheGlyIleThrSerTyrAsp 143
QY 676 TATCACCAAGGAAGTGAACATTTCTCTTTCAAGCCGGTAGTGAATTTATCACGTAA 735
Db 144 PheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheHisCysArg 163
QY 736 GATGAGGGCCACAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGAAACTAGT 795
Db 164 AspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGluIleLysThrGln 183
QY 796 TGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTCGTTT 855
Db 184 CysSerGlyProArgMetAspProLysIleCysProAlaAspProAlaPhePheSerPhe 203
QY 856 ATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAGAGAGACTCACT 915
Db 204 IleAsnAsnSerAspLeuTyrValAlaAsnIleGluThrGlyGluLysArgLeuThr 223
QY 916 TATGTCACATAGCTAGCCACATGGAAGAGATGCCAGATCAGCTGGAGTCCCTACC 975
Db 224 PheCysHisGlnGlyLeuSerAsnValLeuAspProLysIleGluLysArgLeuThr 243
QY 976 TTTGTTCTCCAAAGAAATTTGATGATATCTCGCTATTTGGTGTCTCAAAAGCTGAA 1035
Db 244 PheValIleGlnGluLeuPheAspArgPheThrGlyTyrTrpTyrProThrAlaSer 263
QY 1036 ACACTCCAGCTGGTGTGTTTAAATTTCTATATGAAAGAAATGATGAATCT 1092
Db 264 TrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGluValAspGluSer 283
QY 1093 GAGGTGAAATTTATTCATGTTACATCCCTTATGTTGGAAACAAAGGAGGCGAGATTCATTC 1152
Db 284 GluValGluValIleHisValProSerProAlaLeuGluGluLysThrAspSerTyr 303
QY 1153 CGTTATCTCTAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGATGTCAGAAATATG 1212
Db 304 ArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGln 323
QY 1213 ATTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGCAACTAATTCACCTTTTGAG 1272
Db 324 ThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSer 343
QY 1273 ATTCCTATTGAAGGAGTTGAATATTTATGTCAGAGCTGGATGACTCTCTGAGGGAATAT 1332
Db 344 SerLeuPheProLysValGluTyrIleAlaArgAlaGlyTyrThrArgAspGlyLysTyr 363
QY 1333 GCTTGTGCTCCTACTAGATCCCTCCAGACTCGCTACAGATAGTGTGATCTCACT 1392
Db 364 AlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProPro 383
QY 1393 GAATATTTATCCAGTAGAGATGATGTTATGGAAGGCAGAGACTCAATGATGCTCAGTG 1452
Db 384 AlaLeuPheIleProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaVal 403
QY 1453 CCTGATTTCTGTGAGCCCACTAATTTATCTATGAAGAAACCAACAGACATCTGGATAATTC 1512
Db 404 ProArgAsnValGlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnVal 423

QY 1513 CATGCATCTTTCATGTTTTCCTCCCAAGTCACTC---GAGAGGAAATGAGTTTATTTT 1569
Db 424 HisAspIlePheTyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArg 443
QY 1570 GCTCTGAATGCAAAACAGGTTTCCGTCTATTATACAAAATTAATCATCTATTTTAAAGAA 1629
Db 444 AlaAsnGluCysLysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSer 463
QY 1630 AGCAATATAAACGATCCAGTGGTGGCTGCTGCCAAGTGATTTCAAGTGTCTCTATC 1689
Db 464 GlnGlyTyrAspTyrSerGluProPheSerProGlyGluAspGluPheLysCysProIle 483
QY 1690 AAAGAGGATAGCAATATACAGTGGTGAATGGGAAGTTCTTGGCCGCATGATCTAAT 1749
Db 484 LysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLys 503
QY 1750 ATCCAAGTTGATGAAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTCTTA 1809
Db 504 IleTrpValAsnGluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeu 523
QY 1810 GAGCATCACTCTAGCTAGTACGTTACGTAATCTCGAGAGGTGACAGGCTGACTGAC 1869
Db 524 GluHisHisLeuTyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThr 543
QY 1870 CTTGGCTACTCATCTTCTGTCATCAGTCAGCTGCTGACTTCTTTTAAAGTAAAGTAT 1929
Db 544 ProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyr 563
QY 1930 AGTTAACCAAGAAATCCACACTGTGTCCCTTTACAGACTCACTCAAGTCAAGTCAAGTAC 1989
Db 564 SerSerValSerThrProProCysValHisValTyrLysLeuSerGlyProAspAspAsp 583
QY 1990 CCACTCTGCAAAACAAAGAAATTTTGGCCACCACTTTTGGATTCAGAGGCTCTCTTCT 2049
Db 584 ProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProPro 603
QY 2050 GACTATACTCTCCAGAAATTTCTCTTTTGAATACTACTCGATTTACATTTGATGGG 2109
Db 604 AspTyrValProProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGly 623
QY 2110 ATGCTCTCAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTTCATA 2169
Db 624 MetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheVal 643
QY 2170 TATGTTGCTCTCAGGTCAGTTCGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGC 2229
Db 644 TyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArg 663
QY 2230 TTGAATACCTAGCTCTCTAGTTATGTTGTTGTTAGTATAGACAAACAGGCTCTCTGT 2289
Db 664 LeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCys 683
QY 2290 CACCAGGCTTAAATTTGAAGGCGCTTTAAATATAAATGGTCAATAGAAATGAC 2349
Db 684 GlnArgGlyLeuArgPheGluAlaLeuLysAsnGlnMetGlyGlnValGluIleGlu 703
QY 2350 GATCAGGTGAAGGACTCCATATCTAGCTTCTCGATATGATTTCTATGCTTAGATCTGT 2409
Db 704 AspGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArg 723
QY 2410 GTGGCCTCAGGCTGTCTCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGG 2469
Db 724 ValAlaIleHisGlyTyrPserTyrGlyLysPheLeuSerLeuMetGlyLeuIleHisLys 743
QY 2470 TCAGATATCTTCAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGAT 2529
Db 744 ProGlnValPheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAsp 763
QY 2530 ACAGGATACAGGACGTTATATGTTGGTCACTCCCTGACCAAGTCAAGAGGCTTACTTAA 2589
Db 764 ThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAla 783
QY 2590 GGATCTGTGGCCATGCAAGCAGAGAAAGTTCTCCCTCTGAACAAATCTTACTGTCTTAA 2649

FT	VARSPLIC	788	862	/FTid=VSP_013871. Missing (in isoform 2).	
FT	CONFLICT	369	369	D -> Y (in Ref. 1; BAC37211).	
FT	CONFLICT	546	546	S -> F (in Ref. 2).	
FT	CONFLICT	777	777	Q -> K (in Ref. 1; BAC37211).	
SQ	SEQUENCE	862 AA;	98001 MW;	BID566E824A834E8 CRC64;	
Alignment Scores:					
Pred. No.:		4.18e-201	Length:	862	
Score:		2833.00	Matches:	511	
Percent Similarity:		76.8%	Conservative:	134	
Best Local Similarity:		60.8%	Mismatches:	193	
Query Match:		51.0%	Indels:	2	
DB:		1	Gaps:	2	
US-10-825-632-2 (1-3120) x DPP9_MOUSE (1-862)					
QY	316	TTTTATGTTGAGCGGTATCTCGGAGTCAGCTTAAAGAGCTGCTTCGCGATACACAGAAA	375		
DB	23	PheCysValGlnLysHisSerTrpAspGlyLeuArgSerIleHisGlySerArgLys	42		
QY	376	TATCATGGCTACATGCTAAGCCACCATGATTTTCATGTTCTGTAAGAGGAAATGAT	435		
DB	43	SerSerGlyLeuIleValSerLysAlaProHisAspPheGlnPheValGlnLysProAsp	62		
QY	436	CCAGATGACCTTCATTCAGACAGAACTATTACCTGCCATGCTCGTGAGAACAGAA	495		
DB	63	GluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyrGlySerArgGlu	82		
QY	496	AATACACTGTTTATTCGAAATTCCTCAAACTACTAATAGACAGCAGCTTATATGCTC	555		
DB	83	AsnSerLeuLeuTyrSerGluIleProLysLysValArgLysGluAlaLeuLeuLeuLeu	102		
QY	556	TCTTGGAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATCTGA	615		
DB	103	SerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisGlyValTyrSerArg	122		
QY	616	GAAGAAGAACTATTAAAGAAAGAAAACGCTATCGAAGCAGCTCGGAATTCCTTACGAT	675		
DB	123	GluGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIleThrSerTyrAsp	142		
QY	676	TATCAACCAAGGAAGTGGAACTATCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAA	735		
DB	143	PheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheHisCysArg	162		
QY	736	GATGGAGGCCCAAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGT	795		
DB	163	AspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGluIleLysThrGln	182		
QY	796	TGTCCCAACATACGATCGATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTT	855		
DB	183	CysSerGlyProArgMetAspProLysIleCysProAlaAspProAlaPhePheSerPhe	202		
QY	856	ATACATAGCAACGATTTTGGATATCTAACATCTGTAAACAGAGAACAGGAGACTCACT	915		
DB	203	IleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGluArgArgLeuThr	222		
QY	916	TATGTCCACATGAGCTAGCCCAACATGGAAGAAGATGCCAGATCAGCTGGAAGTCCGTACC	975		
DB	223	PheCysHisGlnGlySerAlaGlyValLeuAspAsnProLysSerAlaGlyValAlaThr	242		
QY	976	TTTGTCTCCAAAGAAATTTGATAGATATTCTGGCTATTGGTGGTCTCCAAAGCTGAA	1035		
DB	243	PheValIleGlnGluPheAspArgPheThrGlyCysTrpTrpCysProThrAlaSer	262		
QY	1036	ACAACCTCCAGTGGTGGT---AAAAATCTTGAATCTTATATGAAGAAAATGATGAATCT	1092		
DB	263	TrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGluValAspGluSer	282		
QY	1093	GAGGTGGAATATTATCATGTTTACATCCCCCTATGTTGGAAACAGAGGAGGAGATTCATTC	1152		
DB	283	GluValGluValIleHisValProSerProAlaLeuGluArgLysThrAspSerTyr	302		

QY	1153	CCTTATCTCTAAACAGGTACAGCAAACTCCTAAAGTCACTTTTAAAGTGTCTAGAAATAATG	1212		
DB	303	ArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluLeuGln	322		
QY	1213	ATTGATGCTGAAGAGGATCATAGATGTCTATAGTAAGCACTAATTCACCTTTTGG	1272		
DB	323	ThrAspHisGlnGlyLysIleValSerSerCysGlnLysGluLeuValGlnProPheSer	342		
QY	1273	ATTCTATTGGAAGAGTTGAATATATTCGCCAGACTGGATGGATCTCTGAGGAAATAT	1332		
DB	343	SerLeuPheProLysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyr	362		
QY	1333	GCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCT	1392		
DB	363	AlaTrpAlaMetPheLeuAspArgProGlnGlnArgLeuGlnLeuValLeuLeuProPro	382		
QY	1393	GAAATATTATCCCATAGTAAGATGTTTATGGAAGCGAGAGACTCATTTGAGTCACTG	1452		
DB	383	AlaLeuPheIleProAlaValGluSerGluAlaGlnArgGlnAlaAlaAlaArgAlaVal	402		
QY	1453	CCTGATTTCTGTGACGCCCACTAATATCTATCGAAGAAAACACAGACATCTGGATAAATATC	1512		
DB	403	ProLysAsnValGlnProPheValIleTyrGluGluValThrAsnValTrpIleAsnVal	422		
QY	1513	CATGACATCTTTCATGTTTTCCTCCCAAGTCAC--GAAGAGGAAATGAGTTTATTTT	1569		
DB	423	HisAspIlePheHisProPheProGlnAlaGluGlyGlnGlnAspPheCysPheLeuArg	442		
QY	1570	GCCTCTGAATGCAAACAGGTTTCCTCATTTTATACAAAATTTACATCTATTTTAAAGCAA	1629		
DB	443	AlaAsnGluCysLysThrGlyPheCysHisLeuTyrArgValThrValGluLeuLysThr	462		
QY	1630	AGCAATATAAAGCATCCAGTGGTGGCTGCTCCAGTGATTTTCAAGTGCCTATC	1689		
DB	463	LysAspTyrAspTrpThrGluProLeuSerProThrGluAspGluPheLysCysProIle	482		
QY	1690	AAAGAGAGATAGCAATACAGTGGTGAATGGAGGTTCTTGGCGGAGCTGATCTAAT	1749		
DB	483	LysGluGluValAlaLeuThrSerGlyGluTrpGluValLeuSerArgHisGlySerLys	502		
QY	1750	ATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCACCAAGACTCCCTTTA	1809		
DB	503	IleTrpValAsnGluGlnThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeu	522		
QY	1810	GAGCATCACCTGTAGTACGTACGTAAATCTCGAGAGGTGACAGGCTGACTGAC	1869		
DB	523	GluHisHisLeuTyrValValSerTyrGluSerAlaGlyGluIleValArgLeuThrThr	542		
QY	1870	CGTGGTCTACTCACATTTCTGCTGATCAGTCAGCTGCTGCTCTTTTATAGTAAGTAT	1929		
DB	543	LeuGlyPheSerHisSerCysSerMetSerGlnSerPheAspMetPheValSerHisTyr	562		
QY	1930	AGTAACCCAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAAGATGAC	1989		
DB	563	SerSerValSerThrProCysValHisValTyrLysLeuSerGlyProAspAspAasp	582		
QY	1990	CCAACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTGGATTCAGCAGGTCCTTCTCT	2049		
DB	583	ProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAlaAsnCysProPro	602		
QY	2050	GACTATCTCTCCAGAAATTTCTCTTTTGAAGCTACTCTGGAATTTTACATTTGATGGG	2109		
DB	603	AspTyrValProGluIlePheHisPheHisThrArgAlaAspValGlnLeuTyrGly	622		
QY	2110	ATGCTCTACAGCTCATGATCTACAGCTCGGAAAGAAATATCCTACTGCTGCTCATATA	2169		
DB	623	MetIleTyrLysProHisThrLeuGlnProGlyArgLysHisProThrValLeuPheVal	642		
QY	2170	TATGGTGGTCTCAGTGGATGGAATATCGGTTTAAAGGATCAAGTATTCCTCCG	2229		
DB	643	TyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArg	662		

```
QY 2230 TTGAATACCTAGCCTCTCTAGTTATGTGGTTATGTGGTTAGTAGACACAGGGATCCTCT 2289
DB |||||||LeuAsnThrLeuAlaSerLeuGlyTyrAlaValValleAspGlyArgGlySerCys 682
QY 2290 CACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAATGGTCAATAGAAATGAC 2349
DB |||||||GlnArgGlyLeuHisPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluLeuGlu 702
QY 2350 GATCAGCTGGAGGACTCCATATCTAGCTTCTCGATATGATTCATTTAGCTAGATCT 2409
DB |||||||AspGlnValGluGlyLeuGlnTyrValAlaGluLysTyrGlyPheLeuAspLeuSerArg 722
QY 2410 GTGGGCATCCAGGCTGGCTTATGAGGATACCTCTCCCTGATGGCATTAAATGACAGAG 2469
DB |||||||ValAlaLeuHisGlyTyrPheSerTyrGlyGlyPheLeuSerLeuMetGlyLeuLeuHisLys 742
QY 2470 TCAGATATCTCAGGCTTCTATCTGCTGGGCCCCAGTCATCTGTGGATCTTCTATGAT 2529
DB |||||||ProGlnValPheLysValAlaLeuAlaGlyAlaProValThrValTrpMetAlaTyrAsp 762
QY 2530 ACAGGATACACGGAAGCTTATATGGGTCAACCTGACACAGATGAACAGGCTTACTTAA 2589
DB |||||||ThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnGlnGlyTyrGluAla 782
QY 2590 GGATCTGTGGCCATGAACAGAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTA 2649
DB |||||||GlySerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuLeu 802
QY 2650 CATGGTTCTCTGATGAGATGTCATCTTGTGACATACCAGTATATCTAGCTGAGTTTAA 2709
DB |||||||HisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeu 822
QY 2710 GTGAGGCTCGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATAAGA 2769
DB |||||||IleArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArg 842
QY 2770 GTTCTCGAATCGGAGAACATTAATGAACTGCACTCTTTTGGCACTACCTTCAAGAAAACCTT 2829
DB |||||||CysArgGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluHisLeu 862

RESULT 5
Q4SBM6_TETNG PRELIMINARY; PRT; 923 AA.
AC Q4SBM6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF14667, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG0020903001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.;
RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
```

```
RN [2]
RP NUCLEOTIDE SEQUENCE.
RL Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014667; CAG01956.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 923
FT NON TER 923
SQ SEQUENCE 923 AA; 105211 MW; 1397023B2004D009 CRC64;

Alignment Scores:
Pred. No.: 7,75e-201 Length: 923
Score: 2829.50 Matches: 524
Percent Similarity: 71.6% Conservative: 138
Best Local Similarity: 56.6% Mismatches: 190
Query Match: 51.0% Indels: 73
DB: 2 Gaps: 7

US-10-825-632-2 (1-3120) x Q4SBM6_TETNG (1-923)
QY 268 GACTGTGAGGAGAAATATTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAG 327
DB |||||||1 AspSerThrGluValValGluMetGluAspValPro-----SerGlnPheValGln 18
QY 328 CGGTATTCTCTGGAGTCAGCTTAAAGAGCTGCTTGGCCGATACACAGAAATATCATGGCTAC 387
DB |||||||19 LysHisSerTrpGluGlyLeuArgAspIleIleHisCysSerArgLysAsnSerGlyIle 38
QY 388 ATGATGGCTAAGCACCACATGATTTTCATGTTTGTGAGAGGATGATCCAGATGGACCT 447
DB |||||||39 IleAlaAsnLysAlaProHisAspPheGlnPheValGlnLysLysAspGluAsnGlyPro 58
QY 448 CATTACAGACAGATCTTATTACCTTGGCCTGCTGCTGAGACAGAGAAATACACTGTTT 507
DB |||||||59 HisSerHisArgLeuTyrTyrLeuGlyMetProTyrGlySerArgGluAsnSerLeuLeu 78
QY 508 TATTCTGAAATCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTTGGAAGCCT 567
DB |||||||79 TyrSerGluLeuProLysLysValArgLysGluAlaLeuLeuValLeuSerTrpLysGln 98
QY 568 CTTTGTGATCTTTTTCAGCAACACATGGAATGATTTCTCGAAGAGAAAGAACTA 627
DB |||||||99 MetLeuAspHisPheGlnAlaThrProHisGlnGlyAlaTyrSerArgGluGluLeu 118
QY 628 TTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687
DB |||||||119 LeuArgGluArgLysArgLeuGlyAlaPheGlyIleThrSerTyrAspTyrHisAlaGln 138
QY 688 AGTGGACATCTTCTGTTTCAAGCCGCTAGTGAATTTATCAGTAAAGATGGAGGGCA 747
DB |||||||139 ThrGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheTyrCysGlnAspGlyGlyGln 158
QY 748 CAGGATTTACGCAACAA---CCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCCAAC 804
DB |||||||159 AsnSerPheIleGlnSerAlaProValLysProValGluIleLysThrGlnCysSerGly 178
QY 805 ATACGGATGATCCAAATTTATCCCGCTGATCCAGACTGGATGCTTTTATCATATAGC 864
DB |||||||179 ThrArgMetAspProLysIleCysProAlaAsnProAspPheIleAlaPheIleAsnAsn 198
QY 865 AACCATATTTGGATATCTAACATCGTAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
DB |||||||199 AsnAspLeuTrpValAlaLeuIleLysThrGlyGluArgArgLeuThrPheCysHis 218
QY 925 AATGAGCTAGCCAACTGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTCTTCTC 984
DB |||||||219 LysGlyAlaAspSerValLysGluAspProLysSerAlaGlyValAlaThrPheValIle 238
QY 985 CAAGAAGATTTGATAGATATCTGCTATTGTTGGTGTGTCCAAAGAGCTGAAACAACTCCC 1044
DB |||||||239 GlnGluGluPheAspArgPheThrGlyTyrTrpTrpSerProSerAlaValGluAspPro 258
```

QY	1045	AGTGGTGTAAATTTCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGAAATTT	1104	QY	1915	TTTATAAGTAAAGTATGTAACACAGAGAATCCACACTGTGTCTCCTTTTACAAGCTATCA	1974
Db	259	AepGlyGlyLysArgValTyrLeuLeuTyrGluGluValAspGluThrGluValGluIle	278	Db	619	PheValSerHisTyrSerSerValCysThrProCysValHisValTyrLysLeuAsn	638
QY	1105	ATTCAATGTTACATCCCTATGTTGGAAACAAGAGGCGCAGATTCCTCGTTATCTCTAA	1164	QY	1975	AGTCTGGAAGATGACCCCAACTTGCAAAACAAGGAATTTTGGCCACCACTTTTGGATTCA	2034
Db	279	IleHieValProSerProAlaLeuLeuGluGluArgLysAlaAspAlaTyrArgTyrProArg	298	Db	639	SerSerGluSerAspProLeuHisIleValProGluPheTrpAlaSerMetMetGluSer	658
QY	1165	ACAGGTACAGCAATCTCTAAAGTCACATTTTAAAGATGTCAGAAATATGATGATCCTGAA	1224	QY	2035	GCAGTCTCTCTTCCTGACTATCTCTCCAGAAATTTCTTTTCCCAAGACTACTCTGGA	2094
Db	299	ThrGlySerLysAsnProGlnAlaThrIleLysLeuValGluIleLysThrAspGlnGln	318	Db	659	SerGlyCysProGlyAspTyrSerProGluIlePheAspPheGlnGlyLysSerGly	678
QY	1225	GGAAGG-----	1230	QY	2095	TTTACATGTTATGGAGATCTCTCAAGCCTCATCATCAGCTCGAGGAAGAATATCTCT	2154
Db	319	GlyArgValSerLeuCysArgLeuLeuCysValPheLeuLeuArgSerAspArgAla	338	Db	679	PheGlnLeuTyrGlyMetValTyrLysProHisSerLeuGlnProGlyArgLysHisPro	698
QY	1231	-----ATCATAGATGTCATAGATAAGCAACTAATTCACACT	1266	QY	2155	ACTGTGCTGTTTCATATATGTTGGTCTCTCAGGTGCAGTGTGTGTAATATCGTTTAAAGGA	2214
Db	339	ValLeuThrCysLeuTyrLeuGlnIleValSerThrGlnAspLysGluLeuAlaValPro	358	Db	699	ThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGly	718
QY	1267	TTTGAGATCTTATTTGAAGAGTTGAATATATGTCAGAGCTGGATGGACTCTCGAGGA	1326	QY	2215	GTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGAC	2274
Db	359	PheThrSerLeuPheProGlyThrGluTyrIleAlaArgValGlyTrpThrSerAspGly	378	Db	719	MetLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAsp	738
QY	1327	AAATAGCTGTGCTCATCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATC	1386	QY	2275	AAACAGGGATCTGTCACCCGAGGGCTTAAATTTTGAAGCGCCTTTAAATATAAAATGGGT	2334
Db	379	LysTyrGlyTrpAlaLeuLeuAspArgSerGlnArgLysLeuGlnLeuValLeuLeu	398	Db	739	GlyArgGlySerCysGlnArgGlyLeuGluPheGluSerAlaLeuLysAsnLysMetGly	758
QY	1387	TCACCTGAATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATGAG	1446	QY	2335	CAAAATAGAATTTGACGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATGATGTTTC	2394
Db	399	ProProAlaPhePheValProValThrAspAspProAlaArgGlnGluSerLeuGlu	418	Db	759	GlnValGluIleGluAspGlnValGluGlyLeuGlnTyrValAlaGluLysPheAsnPhe	778
QY	1447	TCAGTGCCTGATTTCTGTGACCCCACTAATATCTATGAAGAAACAACAGACATCTGGATA	1506	QY	2395	ATTGACTTAGATCGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATG	2454
Db	419	AlaValProAspArgThrGlnProTyrValIleTyrGluGluThrThrAspValTrpIle	438	Db	779	ValAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPheLeuSerLeuMet	798
QY	1507	ATATCCATGACATCTTTCATGTTTTCCTCCCAAGTCCAGAGGAGAAATTTGAGTTTAT	1566	QY	2455	GCATTAATCCAGAGTCAGATATCTCAGGTCCTATGCTGGGGCCCGAGTCACCTCTG	2514
Db	439	AsnValHisAspIlePheTyrProPheValGlnThrAlaGluAspGluPheThrPheIle	458	Db	799	GlyLeuIleGlnArgProAsnValPheLysLeuAlaIleAlaGlyAlaProValThrVal	818
QY	1567	TTTGCTCTGAAATGCAAAACAGGTTTCGTCATTTATACAAAATTACATCTATTTTAAAG	1626	QY	2515	TGATCTTCTTATGATACAGGATACACGGAAGCTTATATGGTACCTGACCCAGCAAGTAA	2574
Db	459	TrpValAsnGluSerLysThrGlyPheSerHisLeuTyrLysIleThrSerValLeuHis	478	Db	819	TrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGln	838
QY	1627	-----GAAAGCAAAATATAACGATCCAGTGTGTGG-----	1656	QY	2575	CAGGGCTATTACTTATAGGATCTGTGCCATGCAAGCAGAAAGTTCCCTCTGACCAAT	2634
Db	479	ProGlyPheHisCysTrpAlaGluAlaTyrHisIleThrGluGlyAspProGlnArgArg	498	Db	839	GlnGlyTyrGluGluGlySerValAlaLeuHisValAspLysLeuProSerGluProAsn	858
QY	1657	CTGCTGCT-----CCAAAGTGAATTTCAAGTCTCCTCTATCAAGAGGAG	1698	QY	2635	CGTTTACTCTTATCATGTTTCTCTGATGAGAATGTCATTTTGACATACCATGATA	2694
Db	499	IleProAlaValSerThrAspValProProGlyAspPheLysCysAlaValLysGluGlu	518	Db	859	ArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPheHisThrAsnPhe	878
QY	1699	ATAGCAATTACAGTGGTGAATGGGAAGTTCTTGGCGGATGATCTAATATCCAGTT	1758	QY	2695	TTACTGAGTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGAG	2754
Db	519	IleThrLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpVal	538	Db	879	LeuValSerGlnIleIleArgAlaGlyLysProTyrGlnLeuGlnValTyrProAsnGlu	898
QY	1759	GATGAAGTCAGAGGCTGTATATTTGAAGCCACCAAGACTCCCTTTTAGAGCATCAC	1818	QY	2755	AGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTATGAATCGATCTTTTGCACAT	2814
Db	539	AsnGluSerSerLysLeuValTyrPheGlnGlyThrArgAspThrProLeuGluHisHie	558	Db	899	ArgHisSerIleArgCysProGluSerGlyGluHisTyrGluIleMetLeuLeuHisPhe	918
QY	1819	CTGTACGTAGTCTAGTTACATAATCTCTGAGAGGTGACAGGCTGACTGACCGTGCTAC	1878	QY	2815	CTTCAAGAAAACCTT	2829
Db	559	LeuTyrValValSerTyrSerProGlyAspValValArgLeuThrLysProGlyPhe	578	Db	919	LeuGlnGlnTyrLeu	923
QY	1879	TCACATTTCTGCTGCATCAGT-----	1899	RESULT 6			
Db	579	SerHisSerCysSerValSerGlnValLysLysSerLeuGlnSerAspTyrPhePheAsn	598	ID	Q6GR22_XENLA	PRELIMINARY;	PRT; 847 AA.
QY	1900	-----CAGCACTGTGACTTC	1914	AC	Q6GR22_XENLA	27, Created)	
Db	599	Tyr-SerSerIleThrLeuProLeuSerLeuSerSerPheIleTrpGlnAsnPheAspPhe	618	DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)	
				DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)	
				DE	MGC81313	protein.	
				GN	Name=MGC81313;		

OS *Xenopus laevis* (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo.
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (May-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071112; AAH71112.1; -; mRNA.
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
 DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR00379; Ser. esters.
 DR Pfam; PF00930; DPPIV_N; I.
 DR Pfam; PF00326; Peptidase_S9; 1.
 SQ SEQUENCE 847 AA; 97484 MW; F515F2609B267BAF CRC64;

Alignment Scores:
 Pred. No.: 8.51e-192 Length: 847
 Score: 2707.50 Matches: 495
 Percent Similarity: 74.3% Conservative: 140
 Best Local Similarity: 57.9% Mismatches: 209
 Query Match: 48.8% Indels: 11
 DB: 2 Gaps: 3

US-10-825-632-2 (1-3120) x Q6GR22_XENLA (1-847)

QY 268 GACTGTGAGGAGATATTGAATCACAGATCGCGCTAAATTGGAGCCTTTTATGTTGAG 327
 Db 3 AspaLaGluGlyArgValAlaSerHis-----PhelysValGlu 15
 QY 328 CGGTATTCCTCGAGTCAGCTTAAAGACTGCTTCCGCGATACCAAGAAATATCATGGCTAC 387
 Db 16 ArgArgSerTrpGluGlyLeuArgAspIleIleHisGlySerArgLysTrpSerGlyIle 35
 QY 388 ATGATGGCTAAGGACCACCATGATGTTTCATGTTTGTGAAGGAGATGATCCAGATGGACCT 447

Db 36 MetValAsnLysAlaProHisAspTyrHisPheGlnSerArgLysGluAspSerGlyThr 55
 QY 448 CATTACAGACAGAAATCTATTACCTTGCATCTGCTGTGAGAACACAGAAAAATACACTGTTT 507
 Db 56 HisSerHisArgLeuTyrPheLeuGlyMetProTyrGlyThrArgGluAsnThrLeuLeu 75
 QY 508 TATTCTGAAATTCCTCAAACTATCAATAGACAGCAGCTTTAAATGCTCTCTTTGGAGCCT 567
 Db 76 TyrSerGluLeuProArgGlnValArgLysGluThrThrLeuLeuLeuSerTrpLysGln 95
 QY 568 CTTTGTGATCTTTTTCAGGCAACACCTGGACTATGAATGATCTCTCGAGAGAAAGAACTA 627
 Db 96 LeuLeuGluAsnPheGlnAlaThrProHisHisGlyValTyrSerArgGluGluGluLeu 115
 QY 628 TTAAGAGAAAGAAAGCATTTGAAACAGTCGGAATTTCTTCTAGCATATACCAAGGA 687
 Db 116 LeuArgGluArgLysArgLeuGlyGlyPheGlyIleThrSerTyrAspPheHisSerLys 135
 QY 688 AGTGGACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCA 747
 Db 136 ThrGlyPhePheLeuPheProAlaAsnAsnSerLeuTyrTyrCysArgAspGlyGlySer 155
 QY 748 CAAGGATTTACGCAACAACTTTAAGSCCAATCTAGTGGAACTAGTTGTCCCAACATA 807
 Db 156 IleSerPheSerAlaSerProValLysProIleGluIleLysThrLeuCysAspGlyPro 175
 QY 808 CGGATGGATCCAAATTTATGCCCGCTGATCCAGACTCGATGCTTTTATACATAGCAAC 867
 Db 176 ArgMetAspAlaLysIleCysProAlaAsnProAspPheIleSerPheIleAsnAsnAsn 195
 QY 868 GATATTTGGATATCTAACTCGTAACACGAGAGAAAGGAGACTCACTTATGTGCACAA 927
 Db 196 AspLeuTrpIleGluAsnIleHisThrMetGluGluLysArgLeuThrTyrCysHisLys 215
 QY 928 GAGCTAGCCAACTGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCAA 987
 Db 216 GlyLeuHisThrValMetAspAspProCysSerAlaGlyValAlaThrPheValIleGln 235
 QY 988 GAAGAATTTGATGATATCTGCTATTTGGTGTGTCTCCAAAGCTGGAACAACACTCCAGT 1047
 Db 236 GluGluPheAspArgPheThrGlyTyrTrpTrpSerProThrValValAspLeuProGly 255
 QY 1048 GGTGTGTAATTTCTAGAAATTTCTATGAGAAATGAGAAATCTGAGTGTGAAATATT 1107
 Db 256 GlyPheSerValLeuArgIleLeuTyrGluGluValAspGluSerAspValGluIle 275
 QY 1108 CATGTTACATCCCTATGTTGGAAACAAGGAGGCGGACTTCATTCGTTATCCTAAAAACA 1167
 Db 276 HisValProSerProAlaLeuGluGluArgLysThrAspAlaTyrArgTyrProArgThr 295
 QY 1168 GGTACAGCAAAATCCTAAAGTCACCTTTTAAGATGTCGAGAAATATGATGATGCTCAAGA 1227
 Db 296 GlySerLysAsnProLysIleThrLeuLysLeuSerGluIleThrThrAsnAsnAspGly 315
 QY 1228 AGGATCATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTCAGATTCCTATTGAAGA 1287
 Db 316 GluIleValAlaAlaGlnAspLysGluLeuValLeuProPheHisSerLeuPheProSer 335
 QY 1288 GTTGAATATATTGCCAGAGCTGGATGAGCTCTCGAGGGAATAATATGCTGGTCCATCTTA 1347
 Db 336 AlaGluTyrIleValArgAlaGlyTrpSerGlnAspGlyLysPheAlaTrpAlaMetLeu 355
 QY 1348 CTAGATCCTCCAGACTCGCTACAGATAGTGTTCATCTCAGCTCACTCAATATTATCCCA 1407
 Db 356 LeuAspArgProGlnCysLeuLeuGlnValLeuLeuProGluLeuPheIlePro 375
 QY 1408 GTAGAGATGATGTTATGGAAGGCGAGAGACTATTGAGTCAGTCGCTGATTCGTGACG 1467
 Db 376 ValCysGluAspGluGluGlnArgIleGlnTyrValLysLysValProGlnHisIleGln 395
 QY 1468 CCACTAATTTATCTATGAGAAACAACAGACATCTCGGATAAATATCCATCAGCATCTTTT 1527

Db	396	ProTyValIleTyGluGluLysSerAspValTrpValAsnValHisAspValValHis	415
QY	1528	GTITTTTCCCAAGATCAGAGAGAAATGAGTTATTTTTCCTCGAATGCAGAAACA	1587
Db	416	LeuLeuArgGlnGluLysGluAsnGluIleArgPheIleCysSerIleGluSerGlnThr	435
QY	1588	GGTTTCCGTCATTTATACAAATTTACATCTTATTTTAAAGAAAGCAATATATAACGATCC	1647
Db	436	GluPheCysHisLeuTyLeuMetThrSerLeuLeuGlnAsnGluAsnArgGluTrpSer	455
QY	1648	AGTGTGGGCTGCTGCCAAGTAT---TTCAAGTGTCTTATCAAGAGAGATAGCA	1704
Db	456	GlnMetThrIleProCysThrValAspMetPheLysCysGlnIleAsnGlnValAla	475
QY	1705	ATTACAGTGTGAATGGAGTCTTTCGGCGCATGGATCTTAATATCAAGTTCATGAA	1764
Db	476	LeuThrSerGlyGluTrpGluValLeuCysArgHisGlyAlaLysValTrpValAsnGlu	495
QY	1765	GTGAGAAGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAC	1824
Db	496	AlaMetGlnLeuValTyPheGlnGlyThrLysAspThrProLeuGluHisLeuTy	515
QY	1825	GTAGTCAGTTACGTAAATCTCGAGAGGTGACAAAGGCTGACTACCGTGGCTACTACAT	1884
Db	516	ValThrSerTyGlnHisProGlyGluValValArgLeuThrGluArgGlyTySerHis	535
QY	1885	TCTTGTCGTCAGTCAGTCAGTCTGACTTCTTTATAGTAAGTATAGTAACCAAGAAAT	1944
Db	536	SerCysThrMetSerProAsnPheAspMetPheValSerGlnTySerSerValSerCys	555
QY	1945	CCACACTGTGTCTCCTTTTACAAAGCTATCAAGTCTCGAAGTACGCCCAACTTGC	2004
Db	556	ProProCysValHisLeuTyArgLeuAsnGly-----AsnProLeuTyGlnHis	572
QY	2005	AAGGATTTTGGGCGACCATTTTGGATTCAGCAGGTCTCTCTCTGACTATATCTCTCA	2064
Db	573	ProGlnPheTrpAlaSerValMetGluAlaLysArgCysProProTyTyValProPro	592
QY	2065	GAATTTTCTCTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGCTT	2124
Db	593	GluIlePheHisPheGlnAlaSerSerGlnValLysLeuTyGlyMetValTyLysPro	612
QY	2125	CATGATCTACAGCTGGAAGAAATATCTACTGCTGCTGCTTCATATATGTCGCTCAG	2184
Db	613	HisAsnLeuValProGlyThrLysHisProThrValLeuPheValTyGlyGlyProGln	632
QY	2185	GTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCC	2244
Db	633	ValGlnLeuValAsnAsnSerPheLysGlyMetLysTyLeuArgLeuAsnThrLeuAla	652
QY	2245	TCTCTAGGTTATGTGTTAGTATAGACACAGGGGATCTCTGTACCAGGGCTTAA	2304
Db	653	HisLeuGlyTyAlaValValIleAspGlyArgGlySerCysHisArgGlyLeuAla	672
QY	2305	TTTGAAGGCGCTTTAAATATAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAGGA	2364
Db	673	PheGluGlyAlaLeuLysAsnLysMetGlyGlnValGluIleGlnAspGlnValGluGly	692
QY	2365	CTCCAATATCTAGCTTCTCGATATGATTTTCATTGATGATGCTGTGGGCTCCAGCGC	2424
Db	693	LeuHisPheValAlaGluArgPheGlyPheValAspLeuAspArgValSerIleHisGly	712
QY	2425	TGTCCTATGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTCAGG	2484
Db	713	TrpSerTyGlyGlyPheLeuSerLeuMetGlyIleIleGlnArgProGluValPheLys	732
QY	2485	GTTGCTATTCCTGGGCGCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGAA	2544
Db	733	ValAlaIleAlaGlyAlaProValThrLeuTrpMetAlaTyAspThrGlyTyThrGlu	752
QY	2545	CGTTATATGGTCCACCTGACAGATGAACAGGGCTATTACTAGGATCTGTGGCCATG	2604
Db	753	ArgTyMetGluThrProGluThrAsnGlnMetGlyTyGluAlaGlySerAlaAlaLeu	772
QY	2605	CMAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTTACTCATGTTTCTCGAT	2664
Db	773	GlnValAspLysLeuProAsnGluProTyArgLeuLeuIleLeuHisGlyPheLeuAsp	792
QY	2665	GAGAAATGTCATTTTGCACATACCATATATTAAGTTTCTAGTGAGGCTCGAAG	2724
Db	793	GluAsnValHisPhePheHisThrAsnPheLeuLeuSerGlnLeuIleArgAlaGlyLys	812
QY	2725	CCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCTCAATCGGA	2784
Db	813	ProTyGlnLeuGlnIleTyProAsnGluArgHisLeuIleArgCysProGluSerGly	832
QY	2785	GAACATTATGAACCTGCATCTTTTGCATCTACCTCAAGAAACCTT	2829
Db	833	GluHisTyGluIleThrLeuLeuHisPheLeuGlnGluHisLeu	847
RESULT 7			
O75273_HUMAN PRELIMINARY; PRT; 508 AA.			
AC	O75273		
DT	01-NOV-1998	(TREMBLrel. 08, Created)	
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	R26984.1	(Fragment)	
OS	Homo sapiens	(Human)	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
NP	[1]		
RP	Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,		
RA	Burkhardt-Schulz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,		
RA	Phan H., Velasco N., Do L., Regala M., Terry A., Georges J., Avila J.,		
RA	Danganan L., Poundstone P., Christensen M., Georges A., Avila J.,		
RA	Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,		
RA	Coeffield J., Duarte S., Lucas S., Bruce R., Thomas P., Ouan G.,		
RA	Kronmiller B., Arellano A., Montgomey M., Ow D., Nolan M., Trong S.,		
RA	Kobayashi A., Olsen A.S., Carrano A.V.,		
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AC005594; AAC33801.1; -; Genomic_DNA.		
DR	MEROPS; S09.019; -.		
DR	Ensembl; ENSG00000142002; Homo sapiens.		
DR	GO; GO:0003824; F:catalytic activity; IEA.		
DR	InterPro; IPR001375; Peptidase_S9.		
DR	InterPro; IPR002469; Peptidase_S9B.		
DR	InterPro; IPR000379; Ser_estrs.		
DR	Pfam; PF00930; DPPIV_N; 1.		
DR	Pfam; PF00326; Peptidase_S9; 1.		
FT	NON_TER	1	
SQ	SEQUENCE	508 AA; 57750 MW; 2F4D6645BE2D2C89 CRC64;	
Alignment Scores:			
Pred. No.:	6.38e-113	Length:	508
Score:	1643.50	Matches:	313
Percent Similarity:	71.8%	Conservative:	77
Best Local Similarity:	57.6%	Mismatches:	108
Query Match:	29.6%	Indels:	45
DB:	2	Gaps:	6
US-10-825-632-2 (1-3120) x O75273_HUMAN (1-508)			
QY	1231	ATCATAGATCTATAGTAAGCACTAATTCACCTTTTGAGATTTCTATTGAGAGGCTT	1290
Db	1	IleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysVal	20
QY	1291	GAATATATTCAGAGCTGGATGGACTCTCTGAGGGAATAATATGCTTGGTCCACTACTA	1350
Db	21	GluTyIleAlaArgAlaGly-----AlaTrpAlaMetPheLeu	33
QY	1351	GATCGCTCCAGACTCCCTACAGATAGTGTGTGATCTCACTGAATTTATCCAGTA	1410

Db	34	AspArgProGlnGlnTrpLeuGlnValLeuLeuProAlaLeuPheIleProSer	53
Qy	1411	GAGATGATGTTATGGAAGCGAGAGACTCATTGAGTCAGCCCTGATTTCTGTACGCCA	1470
Db	54	ThrGluAsnGluGluGlnArgLeuAlaSerAlaAlaValProArgAsnValGlnPro	73
Qy	1471	CTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTT	1530
Db	74	TyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrPro	93
Qy	1531	TTTCCCAAGTACAC---GAAGAGAAATTGAGTTTATTTTGGCTCTGATGCAAAACA	1587
Db	94	PheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThr	113
Qy	1588	GGTTTCGGCTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACGATCC	1647
Db	114	GlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSer	133
Qy	1648	AGTGGTGGCTGCTCTCCAAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGCAATT	1707
Db	134	GluProPheSerProGly-----GluGlyGluGlnSerLeu	145
Qy	1708	ACAGTGGTGAATGGGAAGTCTTGGCCGCGATGATTAATATCAAGTTGATGAAGTC	1767
Db	146	ThrAsnAlaIleTrp-----ValAsnGluGlu	154
Qy	1768	AGAAGGCTGTATATTTGAAGCACCACCAAGACTCCCTTTAGAGCATCACCTGTACGTA	1827
Db	155	ThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeuTyrVal	174
Qy	1828	GTCAGTTACGTAATCTCTGGAGAGTGCACAAAGCTGACTGACCGTGTACTCACATCT	1887
Db	175	ValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSer	194
Qy	1888	TGCTGCTCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAACACGAGAAATCCA	1947
Db	195	CysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrPro	214
Qy	1948	CACGTGTGTCCCTTTACAGCTATCAAGTCTCGAAGATGCCCACTTGCAAAACAAG	2007
Db	215	ProCysValHisValTyrLysLeuSerGlyProAspAspAspProLeuHisLysGlnPro	234
Qy	2008	GAATTTGGCGCCACCATTTTGGATTACAGCAGGTCTCTCTGACTATATCTCCTCAGAA	2067
Db	235	ArgPheTrpAlaSerMetMetGluAlaAla-----Lys	245
Qy	2068	ATTTCCTCTTTGAAGTACTACTGGATTACATTTGATGGATGCTCTACAGGCTCAT	2127
Db	246	IlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHis	265
Qy	2128	GATCTACAGCTGGAAGAAATATCTACTGTGCTCTCATATATATGCTGCTCCTCAGGTG	2187
Db	266	AlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnVal	285
Qy	2188	CAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACCTAGCTCT	2247
Db	286	GlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSer	305
Qy	2248	CTAGGTTATGCTGTGATGATAGACACAGGGGATCCTGTACCGAGGCTTAATTT	2307
Db	306	LeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPhe	325
Qy	2308	GAAGCGCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCAGGTGGAAGACTC	2367
Db	326	GluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeu	345
Qy	2368	CAATATCTAGCTTCTCGATATGATTTTCATTCATCTAGATCGTGGGCATCCACGCTGG	2427
Db	346	GlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrp	365
Qy	2428	TCTATGAGGATACCTCTCCCTGATGCTTAATATGAGAGGTGAGATATCTTCAGGTT	2487
RESULT 8			
Q7QBK1_ANOGA	Q7QBK1_ANOGA PRELIMINARY; PRT; 886 AA.		
ID	Q7QBK1; Q7QBK1; 26, Created		
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	ENSANGP00000016526 (Fragment).		
GN	ORFNames=ENSANG00000014037;		
OS	Anopheles gambiae str. PEST.		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;		
OC	Anophelinae; Anopheles.		
OX	NCBI_TaxID=180454;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=PEST;		
RG	The Anopheles gambiae Sequence Committee;		
RT	"Anopheles gambiae re-annotation,"		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=PEST;		
RG	The Anopheles gambiae Sequence Committee;		
RL	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.		
CC	-!- CAUTION: The sequence shown here is derived from an		
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is		
DR	preliminary data.		
DR	EMBL; AAA01008879; EAA08416.2; -; Genomic_DNA.		
DR	MEROPS; S09.016; -.		
DR	GO; GO:0003824; F:catalytic activity; IEA.		
DR	GO; GO:0008236; F:serine-type peptidase activity; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR01375; Peptidase_S9.		
DR	InterPro; IPR02469; Peptidase_S9B.		
DR	InterPro; IPR00379; Ser. esters.		
DR	Pfam; PF00930; DppIV_N; I.		
DR	Pfam; PF00326; Peptidase_S9; 1.		
FT	NON TER 886		
FT	NON TER 886		
SQ	SEQUENCE 886 AA; 100440 MW; 3C284605CAA57DB4 CRC64;		
Alignment Scores: 1.14e-112 Length: 886			
Pred. No.:			

647 AlaMetValPheLysProHisAsnPheMetLeuGlyValLysTyrProThrValLeuAsn 666
2167 ATATATGGTGGTCTCAGGTGCGATGCGTAATATCGTTTAAAGGAGTCAAGTATTTC 2226
667 ValTyrGlyGlyProGluValGlnThrValSerAsnThrPheLysGlyMetArgGlnLeu 686
2227 CGCTTGAATACCTAGCTCTCTAGGTATGTGTGTAGTCATAGACACACAGGGGATCC 2286
687 ArgMetHisMetLeuAlaSerGlnGlyTyrCysValIleCysValSerArgGlySer 706
2287 TGTCAACGAGGCTTAAATTTGAAGCGGCTTTAAATATAAATGGGTCAATAGAAATT 2346
707 ArgHisArgGlyValGluPheGluSerTyrIleArgCysArgMetGlyThrValGlnLeu 726
2347 GACCATCAGGTGGAAGGACCTCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGAT 2406
727 SerAspGlnValGluValLeuArgIleLeuAlaAspGlnLeuGlyTyrIleAspMetAsp 746
2407 CGGTGGGCAATCCACGCTCGTCTATGAGGAGATACCTCTCCTCATGGCATTAATGAG 2466
747 ArgValAlaIleHisGlyTyrSerTyrGlyGlyTyrLeuSerLeuMetGlyLeuValGln 766
2467 AGTCAATATCTTACGGTGTATGCTGCGGCCCCAGTCACTCTGCGATCTTCTAT 2526
767 TyrProGluIlePheLysValSerIleAlaGlyAlaProValThrSerTyrGluTyrTyr 786
2527 GATACAGATACAGGACGCTTATGCGTCCCTGACCAAGTGAACAGGCGCTATTAC 2586
787 AspThrGlyTyrThrGluArgTyrMetAspLeuProAspSerAsnArgSerGlyTyrAla 806
2587 TTAGCATCTGTGGCCATGCAAGCAGACAAAGTCTCCCTCTGAAACCAATCTTACTGCTC 2646
807 AlaGlySerValLeuAsnTyrIleGlnLysPheProAspGluAspAsnArgLeuLeu 826
2647 TTACATGGTTCCTCGATGAGATGTCATTTGCAATACCATGATATTTACTGAGTTT 2706
827 IleHisGlyLeuIleAspGluAsnValHisPheHisThrSerGlnLeuValSerArg 846
2707 TTAGTGGGCTGGAAGCATATGATTTACAGATCTATCTCCTCAGGAGACACAGATA 2766
847 LeuValArgAlaAsnLysProTyrGlnLeuGlnValTyrProAsnGluArgHisSerLeu 866
2767 AGATTCCTGAATCGGAGACATTAATGACTGCTTTGCACTACCTACCTCAAGAAAC 2826
867 ArgAsnLeuGluAlaSerLysHisTyrGluThrLysLeuSerPheLeuGlnAsnHis 886

RESULT 9
Q9VC20 DROME PRELIMINARY; PRT; 1053 AA.
ID Q9VC20; Q8IH07;
AC 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
DE CG3744-PA, isoform A (CG3744-pc, isoform c) (LUD33755p).
GN ORFNames=CG3744;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazei R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Periera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Furi V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Battencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;

Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 (7)
 RN NUCLEOTIDE SEQUENCE.
 RP STRAINE-Berkeley.

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA Garmez R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

CC -I- INTERACTION:
 CC Q9VZER0:CG12016; NBExp-1; IntAct-EBI-105926, EBI-152819;
 DR EMBL; AB003749; AAF56357.2; -; Genomic DNA.
 DR EMBL; BT001499; AAN71254.1; -; mRNA.

DR InAct; Q9VC20; -;
 DR MEROPS; S09.016; -;
 DR Ensembl; CG3744; Drosophila melanogaster.
 DR FlyBase; FBgn0039240; CG3744.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0003824; F:catalytic activity; IEA.

DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001375; Peptidase_S9.

DR InterPro; IPR002469; Peptidase_S9B.

DR InterPro; IPR00379; Ser esterase.

DR Pfam; PF00930; DPPIV N1.

DR Pfam; PF00326; Peptidase_S9; 1.

SQ SEQUENCE 1053 AA; 118053 MW; C94AA663AB464577 CRC64;

Alignment Scores:

Pred. No.: 2e-111 Length: 1053
 Score: 1624.50 Matches: 393
 Percent Similarity: 53.3% Conservative: 159
 Best Local Similarity: 38.0% Mismatches: 330
 Query Match: 29.3% Indels: 153
 DB: 2 Gaps: 30

US-10-825-632-2 (1-3120) x Q9VC20_DROME (1-1053)

QY 14 TCCGAGGCGAAGCGCGTGTACTGCGCGCGTGTCTTCTTGTAGTGGCGGTGCGCGCTG 73
 DB 72 SerIleThrAenAlaAlaThrSerLeuAlaAen-LeuLeuAaspGlyPheThrAlaAr 91
 QY 74 GGTGTGTCACCGGCGCGCGCGGAGGAGCCACTGCAACCGAGCGGAGGCGGCGGCG 133
 DB 91 gValSerThrAlaAlaAlaAlaAlaAlaAlaAlaAlaThrThrThrGluValValAl 111
 QY 134 CGAGCATGAAGCGCGCGAGCGCGCTCCATAGCG-----CACGTGCGGAGCGTCCGGCG 187
 DB 111 a-----AlaAlaSerValIleThrAenHisLeuSerSerProThrSe 125
 QY 188 GGGG-----CCGGGGGAGGAAATGCAACATGGCAGCAATGGAAACAGAACAGC 241
 DB 125 rGlyThrProProHisGlyLeuAaspValAaspGluGly----- 137
 QY 242 TGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAAATCAGAGTCCGC 301
 DB 138 -----Aap-AaspGluCysAaspCysGluGluAap-----GluAaspAap 151
 QY 302 CTAAATGGAGCCTTTTATGTT-----GAGCGGTATCTCTGGAGTCAAGCTTA 349
 DB 151 rValAaspAenAaspGlyHisLeuAlaAlaProThrProAenHisLeuSerSerProThrA 171
 QY 350 AAAAGCTGCTTCCGATACAGAAATATCATGGTACTATGATGCTAAGGACCAACATG 409
 DB 171 yGlnValValGlnGluValArgGlyLysMetCysAenLeuSerSerMetValProThrA 191
 QY 410 ATTTTCATGTTGTGAAGAGAAATGATCCAGATGAGCTCATTCAGACAGAAATCTTACC 469
 DB 191 snValGlnPhe-----ArgHisLeuSerAaspGly-----ArgAlaArgCysIlePheL 207
 QY 470 TTGCCATGCTGTGGTGAACAGAGAAATACACTGTTTATTCTGAAATTCCTCCAAACTA 529

DB 207 euGlyThrProProGlnSerTrpGluThrThrLeuLeuPheAlaAaspIleAasnLeuThrG 227
 QY 530 TCAATAGAGCAGCAGTCTTAATG----- 552
 DB 227 InSerGluGluGlnGlnLeuLeuValGlnArgLeuGluGlyIleAlaSerAaspGluTrpS 247
 QY 552 ----- 552
 DB 247 erProThrMetAasnAlaGlySerProThrSerSerGlyHisGlnProAlaPheLeuPheA 267
 QY 553 -----CTCTCTGGGAAGCCTCTTTGGATCTT---TTTCAGGCAA 589
 DB 267 snSerLeuProArgProArgLeuProTrpSerProLeuLeuGlnGlnProIleGlnSerS 287
 QY 590 CACTGGACTATGGA-----ATGTATCTTCGAGAAGAAGAACTATTAA 631
 DB 287 erGlyGlySerGlyGlySerGlySerAlaSerProTyrAlaArgGluTyrGlnLeuLeuG 307
 QY 632 GAGAAAGAAAACGCATTGGAAACAGTCGGAATTTGCTTACGATTATTCACCAAGGAAGTG 691
 DB 307 InGluArgGlyArgLeuSerThrTrpGlyIleThrSerTyrGluLeuHisLeuProSerG 327
 QY 692 GAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATTCACGTAAAAGATGGAGGCCACAAG 751
 DB 327 lYlYsLeuValPheProCysPheAasnAaspLeuTyrGlnCysLeuAasp-----ThrG 344
 QY 752 GATTTACGCAACACCTTTTAAGGCCCAATCTAGTGAGAACTAGTTGTCCCAAC---ATAC 808
 DB 344 lYtyrAasnSerGlyLeuLeuPheProThrGlnLeuArgThr---CysProGlnTrpThrA 363
 QY 809 GGATGGATCCAAATATTATGCCCCCTGATCCAGACTGGATTGCTTTTATACATAGCAACG 868
 DB 363 laLeuAaspProGlnIleCysProGlnAenSerAaspMetIleAlaTyrIleSerAaspCysA 383
 QY 869 ATATTGGATATCTAACATCGTACACAGAGAGAGAGAGACTCCTATTATGTGCACAATG 928
 DB 383 spLeuPheValThrHisThrLeuSerGlyHisGluLysArgLeuThrTyrThrSerThrG 403
 QY 929 AGCTAGCCACATGGAAGAGATGCGAGATGCGAGTGGAGTGGTCTTGTCTTCTCAAG 988
 DB 403 lYArgHisSerTyrValAaspAlaLeuSerAlaGlyValProSerTyrValMetGlnG 423
 QY 989 AAGAAATTTGATAGATATCTGCTATGCTGCTGCTCCAAAAGCTGAAACAACTCCCACTG 1048
 DB 423 luGluPheSerArgTyrGlnGlyPheTrpGlnProHis-----SerA 438
 QY 1049 GTGGTAAATCTTAGAATCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATTC 1108
 DB 438 snAaspGlyIleTyrArgIleValTyrGluGluValAaspGluSerGluValSerValTyr 458
 QY 1109 ATGTTACATCCCTATGTTGAAACAAAGGAGGCGAGATTCATTCCTGTTTCTTAAACAG 1168
 DB 458 hrPheProSerSerThrAlaMetHisGlyArgValAaspGluTyrArgPheProArgThrG 478
 QY 1169 GTACAGCAATCTTAAGTCACTTTTAAAGATGTCAGAAATTAATGATTGATGCTGAAGGAA 1228
 DB 478 lYserProAasnAlaLysSerLysLeuLysLeuValGlnPheValLeuAasnGluAlaLeuG 498
 QY 1229 GATCATAGATGTCATAGATAAGAACTAATTAACCTTTTGAGATTCTA-----TTTG 1282
 DB 498 InValSerGluIleAlaIleLysAaspLeu-----ProTyrSerLeuLeuAlaValPheS 516
 QY 1283 AAGGAGTTGAATATATATCCAGAGCTGAGTGGAGTCTCTGAGGGGAAAATATGCTTGGTCCA 1342
 DB 516 erTrpLeuGluTyrIleValArgValGlyTrpThrProAaspAlaLysTyrValTrpValG 536
 QY 1343 TCCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTACCTACCTGAATTTTAA 1402
 DB 536 InGlyLeuAaspArgLysGlnGlnArgLeuAaspValIleLeuIleProLeuAaspAenPhe- 555
 QY 1403 TCCAGTAGAAGATGATGTTATGGAAGGCGAGACTCATTTGAGTCACTGCTCCCTGAT---- 1458

```
Db 556 -----CysGluSerTyrSerSerGlnValSerThrProThrAspSerIleGlyAspHis 574
QY 1459 -----TCTGTGACGCCACTA---ATTATCTATGACAAACAA 1492
Db 574 erTrpArgSerLeuTyrSerArgThrIleThrProLeuGlnValIleTyrThrGluArgS 594
QY 1493 CAGACATCTGGATAAATATCCATGCATCTTTCATGTTTCCCAAAAGTCACCAAGAGG 1552
Db 594 erAspSerTrpIleAsnValHisaspwctLeuHis---PheLeuAspLeuThrGluThrS 613
QY 1553 AAATTTGAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTA 1612
Db 613 erValThrPheLeuTrpAlaSerGlu---GluThrGlyPheArgHisLeuTyrLeuValT 632
QY 1613 CATCTATTTTAAAGGAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCTCA---- 1668
Db 632 hrAlaSerLeuLeuSerGln-----AlaAsnGlyGlnProAspProGlyS 648
QY 1669 -----AGTCATTTCAAGTGTCTCTATCAAGAGG 1696
Db 648 erValGlyAlaGlnProSerPheValAspLeuSerAlaLeuGlnProArgIleLeuAsnL 668
QY 1697 AGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGGCCGCGCATGCATCTAATATCCAAG 1756
Db 668 ysValAlaLeuThrSerGlyGluTrpGluValLeuAlaArg-----AsnLeuTrpV 685
QY 1757 TTGATGAAGTCAGAAAGCTGTGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATC 1816
Db 685 alAspLysAlaAsnLysLeuValTyrPheValGlyLeuArgAspThrProLeuGluLysH 705
QY 1817 ACCTGTACTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGCT 1876
Db 705 isLeuTyrValValSerLeuGluArgProGluHisIleArgLeuLeuThrGluProGlyT 725
QY 1877 ACTCACATTTCTTGC-----TGCATCATCTCAGC 1903
Db 725 yrSerTyrLeuValGluPheAspAspHisPheAsnAspLeuPheProIlePheSerGlnG 745
QY 1904 ACTGTGACTCTTTTAAAGTAAGTATAGTAAC---CAGAAGAATCCACACTGTGTGTCCTC 1960
Db 745 lnCysLysLeuMetLeuLeuValTyrCysAsnIleGlnArgLeuProSerCysLysValM 765
QY 1961 TTTACAGCTATCAAGTCTCGAAGATGCCACCACTTGCAAAACAAAGGAATTTTGGGCCA 2020
Db 765 et-----ArgValAsnGlnThrCysSerAsnGlyGlyValAsnGlyI 779
QY 2021 CCATTTTGGATTACAGAGGT-----CCTCTCTCTGACTATATCTC 2059
Db 779 leGlnIleSerLeuValGlyTyrLeuHisGluGlyGlyLysProGluProGlnTyrCys- 798
QY 2060 CTCACAGAAATTTCTCTTTTGAAGTACTACTCGAATTTACATTTGATGGATGCTCTACA 2119
Db 799 --ProGlnIlePheSerProGlnLeuProSerGlyAspIleValTyrAlaMetValPheL 818
QY 2120 AGCTTCATGATCATACAGCTCGGAAGAAATATCCTACTGTCTGTGTTTCATATATGTTGTC 2179
Db 818 ysProHisAsnPheGluLeuGlyValLysTyrProThrValLeuAsnValTyrGlyGlyP 838
QY 2180 CTCAGGTCAGTGTGGTAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCC 2239
Db 838 roGluValGlnThrValAsnAsnThrPheLysGlyLysHisGlnLeuArgMetHisMetL 858
QY 2240 TAGCTCTCTAGTTTATGTTGTAGTATAGACAAACAGGGGATCTGTCTACCGAGGCC 2299
Db 858 euAlaGlnGlyTyrCysValIleCysIleAspSerArgGlySerArgHisArgGlyL 878
QY 2300 TTAATTTGAGGGCCCTTAAATATATAAATGGTCAATAGATTAATTCAGTCATCAGTGG 2359
Db 878 ysArgPheGluSerHisIleArgGlyArgMetGlyGlnValGluLeuThrAspGlnValA 898
QY 2360 AAGGACTCCAATATCTAGTCTCTCGATATGATTTTCATTCATCTAGATCTGTGGGCGATCC 2419
Db 898 spAlaLeuArgSerLeuSerAspGlnLeuGlyTyrIleAspMetAspArgValAlaIleH 918
```

```
2420 ACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTCAGATATCT 2479
Db 918 isGlyTrpSerTyrGlyGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProLysIleP 938
QY 2480 TCAGGGTGTCTATGTCTGGGGCCCCAGTCACCTCTCTGTGGATCTTCTATGATACAGATACA 2539
Db 938 heLysValAlaIleAlaGlyAlaProValThrAsnTrpGluTyrTyrAspThrGlyTyrT 958
QY 2540 CGGAACGTTATATGGTTCACCTCCCTGACCAAGATGAACAGGGCTATTACTTAGATCTGTGG 2599
Db 958 hrGluArgTyrMetAspMetProGlnAsnAsnGluAlaGlyTyrSerAlaGlySerValL 978
QY 2600 CCATGCAAGCAGAAAGTCTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCC 2659
Db 978 euGluTyrValAsnSerPheProGluGluAspLysArgLeuLeuIleHisGlyLeuI 998
QY 2660 TGGATGAAGATGTCCATTTTGCATACCATCATATATTTACTGAGTGTCTTTTAGTGAGGCTG 2719
Db 998 leAspGluAsnValHisPheCysHisThrSerArgLeuIleSerAlaLeuAsnLysAlaA 1018
QY 2720 GAAAGCCATATGATTTTACAGATCTCTCTCAGAGAGACACAGATCAATAGAGTTCCTGAAT 2779
Db 1018 snLysProTyrGluValHisLeuPheProGluGluArgHisSerLeuArgAsnLeuGluS 1038
QY 2780 CGGAGAACATTAATGAACATCTCTTTTGCATCTACTCTCAA 2820
Db 1038 erAsnLysAsnTyrGluThrLysLeuLeuSerPheLeuGln 1051

RESULT 10
Q9VC19 DROME PRELIMINARY; PRT; 1113 AA.
ID Q9VC19
AC Q9VC19
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE CG3744-PB, isoform B.
GN Name=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
```


Db	447	rHisThrLeuSerGlyHisGluValArgLeuThrThrSerThrGlyArgHisSerTy	467	Db	789	IgluPheAspAspHisPheAsnAspLeuPheProIlePheSerGlnGlnCysLysLeuMe	809
Qy	942	GGAAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCAGAGAAATTTGATAG	1001	Qy	1917	TATAAGTAAGTATAGTAAC---CAGAAGAATCCACACTGTGTGTCTTACCAAGCTATC	1973
Db	467	rValAspAspAlaLeuSerAlaGlyValProSerTyrValMetGlnGluPheSerAr	487	Db	809	tLeuLeuValTyrCysAsnIleGlnArgLeuProSerCysLysValMet-----	825
Qy	1002	ATATCTGGCTATTGGTGTGTCCAAAGAGCTGAACAACTCCAGTGGTGTAAATCT	1061	Qy	1974	AAGTCTCGAAGATGCCAACCTTCGCAAAACAAGAAATTTGGGCGCCACCATTTGGATTC	2033
Db	487	gTyrGlnGlyPheTrpGlnProHis-----SerAsnAspGlyIleTy	502	Db	826	-----ArgValAsnGlnThrCysSerAsnGlyGlyValAsnGlyIleGlnIleSerLe	843
Qy	1062	TAGAATCTATATGAAGAAATGATGATCTCAGGTGGAAATTTATTCATGTTACATCCC	1121	Qy	2034	AGCAGGT-----CCTCTCTCAGCTATATCTCTCCAGAAATTTT	2072
Db	502	rArgIleValTyrGluGluValAspGluSerGluValSerValTyrThrPheProSerSe	522	Db	843	uValGlyTyrLeuHisGluGlyGlyLeuProGlnTyrCys---ProGlnIlePh	862
Qy	1122	TATGTTGGAAACAAGAGCGGAGATTCATTCCTGTTATCTTAAACAGGTACAGCAATCC	1181	Qy	2073	CTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAGCCTCATGATCT	2132
Db	522	rThrAlaMetHisGlyValAspGluTyrArgPheProArgThrGlySerProAsnAl	542	Db	862	eSerProGlnLeuProSerGlyAspIleValTyrAlaMetValPheLysProHisAsnPh	882
Qy	1182	TAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGCTGAAGAGGATCATGATGT	1241	Qy	2133	ACAGCTCGGAAAGAAATATCTTACTGCTGCTTATATATATATATATATATATATAT	2192
Db	542	alySerLysLeuLysLeuValGlnPheValLeuAsnGluAlaLeuGlnValSerGluI	562	Db	882	eGluLeuGlyValLysTyrProThrValLeuAsnValTyrGlyGlyProGluValGlnTh	902
Qy	1242	CATAGATAAGAACTAATCAACCTTTTGAGATTTCTA-----TTTGAAGAGATTGAATA	1295	Qy	2193	GGTGAATATATGCGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCTTAGCTCTTAGG	2252
Db	562	eAlaIleLysAspLeu-----ProTyrSerLeuLeuAlaValPheSerTrpLeuGluTy	580	Db	902	rValAsnAsnThrPheLysGlyLysHisGlnLeuArgMetHisMetLeuAlaAlaGlnGl	922
Qy	1296	TATTCGCCAGCTGATGATGCTGAGGAGAAATATGCTTGGTCCATCTCTACATGATCG	1355	Qy	2253	TTATGTTGGTTGATGATAGCAACAGGGGATCTCTGTCCACGAGGGCTTAAATTTGAAGG	2312
Db	580	rIleValArgValGlyTrpThrProAspAlaLysTyrValTrpValGlnGlyLeuAspAr	600	Db	922	yTyrCysValIleCysAlleAspSerArgGlySerArgHisArgGlyLysArgPheGluSe	942
Qy	1356	CTCCGACCTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTCCCGATAGAGA	1415	Qy	2313	CGCTTTAAATATAAATGGGTCAATAGAAATTCACCATCAGGTGGGAGGACCTCCAATA	2372
Db	600	glysGlnGlnArgLeuAspValIleLeuLeuProLeuAspAsnPhe-----CysGluSe	618	Db	942	rHisIleArgGlyArgMetGlyGlnValGluLeuThrAspGlnValAspAlaLeuArgse	962
Qy	1416	TGATGTTATGGAAGGCAGAGACTCATTCAGTCAGTGCCTGAT-----	1458	Qy	2373	CTAGCTTCTCGATATGATTTTCATTCAGTTCAGTCTGGGCGATCCACGGCTGTCTCTA	2433
Db	618	rTyrSerSerGlnValSerThrProThrAspSerIleGlyAspHisSerTrpArgSerLe	638	Db	962	rLeuSerAspGlnLeuGlyTyrIleAspMetAspArgValAlaIleHisGlyTrpSerTy	982
Qy	1459	-----TCTGTGACGCCACTA---ATTATCTATGAAGAAACAACAGATCTGGAT	1505	Qy	2433	TGAGAGATACCTCTCCTCATGATTAATGATGAGAGGTGATATCTTCAGGGTTGCTAT	2499
Db	638	uTyrSerArgThrIleThrProLeuGlnValIleTyrThrGluArgSerAspSerTrpI	658	Db	982	rGlyGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProLysIlePheLysValAlaI	1000
Qy	1506	AAATATCATGATCTTTTATGTTTCCCAAGTACAGAGAGAAATGATGATTTAT	1565	Qy	2493	TGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGATACAGGACGTTATAT	2555
Db	658	eAsnValHisAspMetLeuHis---PheLeuAspLeuThrGluThrSerValThrPheLe	677	Db	1002	eAlaGlyAlaProValThrAsnTrpGluTyrTyrAspThrGlyTyrThrGluArgTyrMe	1022
Qy	1566	TTTTGCTCTGAATGCAAAACAGGTTTCGTCATTTATACAAATTTACATCTATTAA	1625	Qy	2553	GGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTGGATCTGTGGCCATGCAAGCAGA	2611
Db	677	uTrpAlaSerGlu---GluThrGlyPheArgHisLeuTyrLeuValThrAlaSerLeuLe	696	Db	1022	tAspMetProGlnAsnAsnGluAlaGlyTyrSerAlaGlySerValLeuGluTyrValas	1041
Qy	1626	GGAAGCAATATAACGATCCAGTGTGGGCTGCTGCTCCA-----	1668	Qy	2613	AAAGTTCCCTCTGAACCAAAATCGTTTACTGCTTCTTACATGTTCTCGATGCAAGTGT	2677
Db	696	uLeuSerGln-----AlaAsnGlyGlnProAspProGlySerValGlyAlaGl	712	Db	1042	nSerPheProGlnAsnAsnGluAlaGlyTyrSerAlaGlySerValLeuGluTyrValas	1061
Qy	1669	-----AGTGATTTCAAGTCTCCTATCAAAAGAGAGATAGCAATTAAC	1709	Qy	2673	CGATTTTGACATACACAGATATATTACTGAGTGTTTTGTAGTGGGTGGAAGCCATATAGA	2733
Db	712	nProSerPheValAspLeuSerAlaLeuGlnProArgIleLeuAsnLysValAlaLeuTh	732	Db	1062	lHisPheCysHisThrSerArgLeuLeuSerAlaLeuAsnLysAlaAsnLysProTyrGl	1081
Qy	1710	CAGTGGTCAATTTGGAGCTTCTTGGCGGCAATGATCTTAATATCAAGTTGATGAAGTCAG	1769	Qy	2733	TTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCTGATCGGAGAGACATTA	2799
Db	732	rSerGlyGluTrpGluValLeuAlaArg-----AsnLeuTrpValAspLysAlaAs	749	Db	1082	uValHisLeuPheProGluGluArgHisSerLeuArgAsnLeuLeuGluSerAsnLysAsnTy	1101
Qy	1770	AAGCTGGTATATTTGAAGCCCAAGAAAGTCCCTTTAGACATCACCTGTACGTAGT	1829	Qy	2793	TGAAGTGCATCTTTTGCATCTACCTTCAA 2820	
Db	749	nLysLeuValTyrPheValGlyLeuArgAspThrProLeuGluLysHisLeuTyrValVa	769	Db	1102	rGluThrLysLeuLeuSerPheLeuGln 1111	
Qy	1830	CAGTGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTGCTCTCATCTTCTG	1889	Qy	2820	-----	
Db	769	lSerLeuGluArgProGluHisIleArgLeuLeuThrGluProGlyTyrSerTyrLeuVa	789	Db			
Qy	1890	C-----TGCAATCAGTCAGCTGTGACTTCTT	1916	Qy			


```
QY 2062 -----CCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTGATGGG 2109
Db 481 ThrGlnTyrAsnProSerIleHisSerProGlnIleSerSerGlyAspValLeuTyrAla 500
QY 2110 ATGCTCTACAGCCTCATGATCATACAGCCTGGAAAGAAATATCCCTACTGTGCTGTTCATA 2169
Db 501 MetValPheLysProHisAsnPheMetLeuGlyValLysTyrProThrValLeuAsnVal 520
QY 2170 TATGGTGGTCTCTCAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGCG 2229
Db 521 TyrGlyGlyProGluValGlnThrValSerAsnThrPheLysGlyMetArgGlnLeuArg 540
QY 2230 TTGAATACCTAGCCTCTCTAGTTATGCTGTGTAGTATGATAGACACACAGGGGATCCTGT 2289
Db 541 MethisMetLeuAlaSerGlnGlyTyrCysValIleCysValAspSerArgGlySerArg 560
QY 2290 CACGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATAGAAATTCAC 2349
Db 561 HisArgGlyValGluPheGluSerTyrIleArgCysArgMetGlyThrValGluLeuSer 580
QY 2350 GATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCGT 2409
Db 581 AspGlnValGluValLeuArgIleLeuAlaAspGlnLeuGlyTyrIleAspMetAspArg 600
QY 2410 GTGGGCATCCAGCGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGG 2469
Db 601 ValAlaIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyr 620
QY 2470 TCAGATATCTTCAGGGTGTCTATTGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGAT 2529
Db 621 ProGluIlePheLysValSerIleAlaGlyAlaProValThrSerTrpGluTyrTyrAsp 640
QY 2530 ACAGATACACGAGACGTTATATGGGTCCACCTGACACAGATGACAGACAGGCGCTTACTTA 2589
Db 641 ThrGlyTyrThrGluArgTyrMetAspLeuProAspSerAsnArgSerGlyTyrAlaAla 660
QY 2590 GGATCTGGGCCATGCACAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTA 2649
Db 661 GlySerValLeuAsnTyrIleGlnLysPheProAspGluAspAsnArgLeuLeuIle 680
QY 2650 CATGGTCTTCGGATGAGTATGCTATTTGCACATACAGTATATTTACTGAGTTTCTTA 2709
Db 681 HisGlyLeuIleAspGluAsnValHisPheHisThrSerGlnLeuValSerArgLeu 700
QY 2710 GTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACACAGCATAAGA 2769
Db 701 ValArgAlaAsnLysProTyrGlnLeuGlnValTyrProAsnGluArgHisSerLeuArg 720
QY 2770 GTTCTCGAATCGGAGAACATATTGAATCGATCTTTTGGCACTACCTTCAAGAAACCTTT 2829
Db 721 AsnLeuGluAlaSerLysHisTyrGluThrLysLeuSerPheLeuGlnAsnHisLeu 740

RESULT 12
ID Q7PTT8 ANOGA PRELIMINARY; PRT; 621 AA.
AC Q7PTT8-
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ENSANGP0000015447 (Fragment).
GN ORFNames=ENSANGP0000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008287; EAA03335.3; -, Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DDPV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1 621
FT NON_TER 621 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;

Alignment Scores:
Pred. No.: 5,68e-89 Length: 621
Score: 1321.00 Matches: 279
Percent Similarity: 58.6% Conservative: 99
Best Local Similarity: 43.3% Mismatches: 187
Query Match: 23.8% Indels: 80
DB: 2 Gaps: 13

US-10-825-632-2 (1-3120) x Q7PTT8 ANOGA (1-621)
QY 1054 AMAATCTTGAATCTTATATGAGAAATGATGATCTGAGTGGAAATTTATTCATGTT 1113
Db 3 GluValTyrArgIleValTyrGluValAspGluSerAspValSerLeuTyrThrPhe 22
QY 1114 ACATCCCTATGTGTGGAACAAGAGGCGACATTCATTCCTGTTATCTTAAACAGGTACA 1173
Db 23 ProSerSerGlnSerAlaGlyArgAspTyrGluGluTyrArgPheProArgAlaGlyThr 42
QY 1174 GCAATCTCTAAGTCACTTTTAAGATGTCAGAAATATGATGATCTGCTGAAGGAGGATC 1233
Db 43 ProAsnAlaLysSerLysLeuLysValGlnPheArgLeuSerGluAsnLeuArgIle 62
QY 1234 ATGATGTCATAGATAGGAACTAATTCACCTTTTGAGATTCATTTTGAAGGAGTTCAA 1293
Db 63 ThrAspValCysIleLysGluLeuGlnCysProLeuThrPheAlaPheProTrpLeuGlu 82
QY 1294 TATATTCCAGAGCTCGATGGACTCTGAGGAGAAATATGCTTGCTCATCTACTAGAT 1353
Db 83 TyrIleValArgValGlyTrpThrProAspSerArgTyrValTrpAlaGlnLeuAsp 102
QY 1354 CGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTTATCCAGTAGAA 1413
Db 103 ArgProGlnGlnArgLeuLeuValLeu-----LeuProVal--- 115
QY 1414 GATGATGTTTGAAGAGGAGGAGACTCATATTGAGTCAGTGCCT----- 1455
Db 116 AspAsnPheCysGluIleTyrSerSerSerSerSerSerLeuProProArgLysSerSer 135
QY 1456 -----GATTCGTGAGCGCCACTA---ATTATCTATCAAGAA 1488
Db 136 GlyTrpArgSerProLeuAspLysSerThrThrArgProLeuGlnValIleTyrThrGlu 155
QY 1489 ACAACAGACATCTGGATAATATCCATGACATCTTTCATGTTTTCCTTCCCAAGTCACGAA 1548
Db 156 ThrSerSerSerTrpValAsnValHisAspValLeuGln---PheValGluLeuSerGlu 174
QY 1549 GAGGAAATTCAGTTTATTTTTCCTGCTGAAATGCAAAACAGAGTTTCCGTCATTATACAAA 1608
Db 175 GlnGluValThrPheLeuTrpAlaSerGlu-----GluSerGlyPheArgHisLeuTyrLeu 193
QY 1609 ATTACATCTATTTTAAAGGAGAAACAAATATATAACGATCCAGTGGTGGGCTGCCTCCA 1668
```

```
Db 194 ValThrSerSerLeuSerProAsnGluVal-----SerSerGlyGlyValalaGlyAla--- 210
Qy 1669 AGTCATTTCAAGTGCTCT-----ATCAA 1692
Db 211 ThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuValAlaAargIleVal 230
Qy 1693 GAGGAGATAGCATACCAAGTGTGAATGGGAAGTCTTTGGCGGCATCGATCTAATATC 1752
Db 231 GlnLysValThrLeuThrGlyGlyAspTrpGluValLeuGlyArg-----AsnVal 247
Qy 1753 CAAGTTGATGAAGTCAAGAGCTGGTATTTTGAAGGCCCAAGACTCCCTTTAGAG 1812
Db 248 TrpTyrAspArgValaArgGlnLeuValTyrPheMetGlyLeuAargIleThrProLeuGlu 267
Qy 1813 CATCACCTGACGTAGTACGTAAATCTCGGAGAGGTGACAAGCTGACCTGACCGGT 1872
Db 268 LysHisLeuTyrValValSerLeuAlaGlnProAsnGlnLeuAargLeuThrMetPro 287
Qy 1873 GGCTACTCACATCTTCTGCTGCATCAGTCAGCTGTGACTTCTTTAAGTAAGTATAGT 1932
Db 288 GlyTyrSerPheThrValGluPheAsnAspAspCysThrLeuPheLeuGlnThrTyrCys 307
Qy 1933 AACCCAGGAAGATCCACACTGTGTGCTCCCTTTACAAGCTA-----TCAAGTCCCTGAA 1983
Db 308 AsnIleSerThrLeuProSerTrpGluLeuValaValaGileAlaHisAspSerAsnThrAla 327
Qy 1984 GATGACCCAACTTGCAAAACAAGGAATTTGGCCACCATTTCGGATTCAGCAGGTCT 2043
Db 328 AsnGlyAsnGlyCys-----SerHisGlyPro 336
Qy 2044 CTTCTGACTATCTCTCT----- 2061
Db 337 ThrProProThrProThrIleAspAlaLeuAargLeuCysSerValGlyTyrLeuThrGlu 356
Qy 2062 -----CCGAAATTTCTCTTTTGAAGTACTACT 2091
Db 357 GlyGlyProSerGluAsnThrGlnTyrAsnProSerIleHisSerProGlnIleSerSer 376
Qy 2092 GGATTTACATGTATGGGATGCTCTACAGCTCATGATCTACAGCTCGGAAAGAAATAT 2151
Db 377 GlyAspValLeuTyrAlaMetValPheLysProHisAsnPheMetLeuGlyValLysTyr 396
Qy 2152 CCTACTGTGCTGTTCATATATATGTGTGCTCTCAGGTGCAGTTCGGTGAATAATCGGTTAAA 2211
Db 397 ProThrValLeuAsnValTyrGlyGlyProGluValGlnThrValSerAsnThrPheLys 416
Qy 2212 GGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTATGTGTGTGTAGTGATA 2271
Db 417 GlyMetArgGlnLeuAargMetHisMetLeuAlaSerGlnGlyTyrCysValIleCysVal 436
Qy 2272 GACAAAGGGGATCCTGTCCAGGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATG 2331
Db 437 AspSerArgGlySerArgHisArgGlyValGluPheGluSerTyrIleAargArgMet 456
Qy 2332 GGTCAAAATAGAAATTTGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGAT 2391
Db 457 GlyThrValGluLeuSerAspGlnValGluValLeuAargIleLeuAlaAspGlnLeuGly 476
Qy 2392 TTCATTGACTTAGATCTGTGGGCTCCAGCTGCTCTATGAGGAGTACTCTCCCTG 2451
Db 477 TyrIleAspMetAspArgValAlaIleHisGlyTyrSerTyrGlyTyrLeuSerLeu 496
Qy 2452 ATGGCATATATGAGAGGTGATATCTCAGGTTGCTATGCTGGGCGCCGCTACT 2511
Db 497 MetGlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProValThr 516
Qy 2512 CTGTGGATCTCTATGATACAGGATACAGGACGCTTATATGGGTCACTCCCTGACAGAT 2571
Db 517 SerTrpGluTyrTyrAspThrGlyTyrThrGluArgTyrMetAsnLeuProAsnSerAsn 536
Qy 2572 GAACAGGGCTATTAATTAGATCTGTGGCCATCCAGCAGAAAGTTCCCTCTCGAACCA 2631
```

```
Db 537 ArgSerGlyTyrAlaAlaGlySerValLeuAsnTyrIleGlnLysPheProAspGluAsp 556
Qy 2632 RATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATCTCCATTTTGCACATACCAGT 2691
Db 557 AsnAargLeuLeuIleIleHisGlyLeuIleAspGluAsnValHisPheHisThrSer 576
Qy 2692 ATATTACTAGCTTTTGTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAG 2751
Db 577 GlnLeuValSerArgLeuValaArgAlaAsnLysProTyrGlnLeuGlnValTyrProAsn 596
Qy 2752 GAGAGACACAGATGAAGTTCCTGAATCGGAGAACATTAATGACATGTCATCTTTGCAC 2811
Db 597 GluAargHisSerLeuAargAsnLeuGluAlaSerLysHisTyrGluThrLysLeuLeuSer 616
Qy 2812 TACCTTCAAGAAAC 2826
Db 617 PheLeuGlnAsnHis 621
```

RESULT 13

Q5TXJ2 ANOGA PRELIMINARY; PRT; 557 AA.

AC Q5TXJ2;

DT 01-FEB-2005 (TReMBLrel. 29, Created)

DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)

DE ENSANGP0000029249 (Fragment)

GN ORFNames=ENSANG0000012958;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;

OC Anophelinae; Anopheles.

OX NCBI_TaxID=180454;

RN (1)

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RT "Anopheles gambiae re-annotation."

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RN (2)

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.

RL -!- CAUTION: the sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; AAA01008287; EAL42106.1; -; Genomic_DNA.

DR GO; GO:0003824; F: catalytic activity; IEA.

DR GO; GO:0008236; F: serine-type peptidase activity; IEA.

DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.

DR InterPro; IPR001375; Peptidase_S9.

DR InterPro; IPR002469; Peptidase_S9B.

DR InterPro; IPR000379; Ser. esters.

DR Pfam; PF00930; DPPIV_N; I.

DR Pfam; PF00326; Peptidase_S9; 1.

FT NON_TER 1 557

FT NON_TER 557 557

SQ SEQUENCE 557 AA; 62685 MW; 08AAA0944ACD0808 CRC64;

Alignment Scores:

Pred. No.:	9,648-76	Length:	557
Score:	1142.50	Matches:	253
Percent Similarity:	57.7%	Conservative:	84
Best Local Similarity:	43.3%	Mismatches:	198
Query Match:	20.6%	Indels:	49
DB:	2	Gaps:	12

US-10-825-632-2 (1-3120) x Q5TXJ2_ANOGA (1-557)

Qy 1063 AGAATCTATATGAGAAATGATCAATCTGAGTGGAAATATTATTCATGTTACATCCCT 1122

Db 1 ArgIleValTyrGluGluValAspGluSerAspValSerLeuTyrThrPheProSerSer 20

```
QY 1123 ATGTTGGAACAAGAGGCGCAGATTCAATCCGTTATCTCTAAACAGGTACAGCAAAATCTCT 1182
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
21  GlnSerAlaGlyArgAspTyrGluGluTyrArgPheProArgAlaGlyThrProAsnAla 40
QY 1183 AAAGTCACATTTAAGATGTCAGAAATATGATGCTGCTGAGGAAGGATCATAGATGTC 1242
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
41  LysSerLysLeuLysLeuValGlnPheArgLeuSerGluAsnLeuArgIleThrAspVal 60
QY 1243 ATAGATAAGGAACATAATCAACCTTTTGAGATTCATTATGAGGAGTGAATATATTGCC 1302
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
61  CysIleLysGluLeuGlnCysProLeuThrPheAlaPheProTrpLeuGluTyrIleVal 80
QY 1303 AGAGCTGGATGACCTCCTGAGGAAAATATGCTTGGTCCATCTCTACTAGATCCCTCCAG 1362
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
81  ArgValGlyTrpThrProAspSerArgTyrValTrpAlaGlnLeuLeuAspArgProGln 100
QY 1363 ACTCGCTCACAGATAGTGTGATCTCACCTCAATATTATTCCAGTAGAAGATGATGTT 1422
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
101  GlnArgLeuGluLeuValLeuLeuProValAspAsnPheCysSerThrAla-----Ala 118
QY 1423 ATGGAAGGACAGACATCATTTGAGTCAGTGCCTGATCTGTGACGCCACATAATTATCTAT 1482
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
119  ArgArgArgProThrAlaGlyArgProProGlyAlaMetGlyProGlyIleTyr 138
QY 1483 GAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTCCTCCCAAAGT 1542
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
139  ThrGluThrSerSerTrpValAsnValHisAspValLeuGln---PheValGluLeu 157
QY 1543 CACGAAGAGAAATGAGTTATTTTTCCTCTGANTGCCAAACAGAGTTTCCGTCAATTTA 1602
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
158  SerGluGlnGluVal-----ThrThrAsnSerHisPro 168
QY 1603 TACAAATTTACATCTATTTTAAAGCAAGCAAAATAT-----AAACGATCCAGTGTGGG 1656
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
169  ThrAlaHisProLysGlnGlnArgGluGlyAlaHisGlyThrGluValSerSerGlyGly 188
QY 1657 CTGCCTGCTCCAAAGTATTCAAGTGTCT-----1686
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
189  AlaGlyAla---ThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuVal 207
QY 1687 -----ATCAAGAGGAGATAGCAATACCAGTGGTGAATGGGAAGTCTTGGCCGCGCAT 1740
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
208  AlaArgIleValGlnLysValThrLeuThrGlyGlyAspTrpGluValLeuGlyArg---226
QY 1741 GGATCTAATATCCAAAGTTGATGAGTCAGAGCGTGGTATATTTTGAAGCAGCAACAAGAC 1800
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
227  -----AsnValTrpTyrAspArgValArgGlnLeuValTyrPheMetGlyLeuArgGlu 244
QY 1801 TCCCTTTTAGACATCACCCTGACGTAGTCAGTAAATCCTCGAGAGGTGACACAGG 1860
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
245  ThrProLeuGluLysHisLeuTyrValValSerLeuAlaGlnProAsnGlnLeuArgLeu 264
QY 1861 CTGACGTGACGGTGGTACTACATCTTGTGCTGCATCAGTCAGCAGTGTGACTTCTTT---1917
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
265  LeuThrMetProGlyTyrSerPheThrVal-----GluPheAsnAspLeuValArg 281
QY 1918 ATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTTTTACAAGTATCAAGT 1977
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
282  IleAlaHisAspSerAsnThrAlaAsnGlyAsnGlyCysSer-----HisGly 297
QY 1978 CTGGAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCACCACTTTTGGATTCAGCA 2037
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
298  ProThrProProThr-----ProIleAspAlaLeuArgLeuCysSerVal 313
QY 2038 GGTCTCTTCTCCTGATATATCTCT-----CCAGAAATTTTCTCT 2076
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
314  GlyTyrLeuThrGluGlyGlyProSerGluAsnThrGlnTyrAsnProSerIleHisSer 333
QY 2077 TTTGAAGTACTACTGGATTTACATTGATGGATGCTCTACAGCCTCATGATCTACAG 2136
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
334  ProGlnIleSerSerGlyAspValLeuTyrAlaMetValPheLysProHisAsnPheMet 353
QY 2137 CTTGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCTCAGGTGCATGTTGGT 2196
```

```
Db 354 LeuGlyValLysTyrProThrValLeuAsnValTyrGlyGlyProGluValGlnThrVal 373
QY 2197 AATAATCGGTTTAAAGAGGTCAAGTATTTCGCTTGAATACCCCTAGCTCTCTAGGTTAT 2256
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
374  SerAsnThrPheLysGlyMetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyr 393
QY 2257 GTGGTTGTATGATAGACAAACAGGGGATCTGTCTACCGAGGGCTTAAATTTGAAGGCGCC 2316
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
394  CysValIleCysValAspSerArgGlySerArgHisArgGlyValGluPheGluSerTyr 413
QY 2317 TTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATAGGTGGAGGACTCCAATATCTA 2376
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
414  IleArgArgArgMetGlyThrValGluLeuSerAspGlnValGluValLeuArgIleLeu 433
QY 2377 GCTTCTCGATATGATTTTCACTTACCTAGATCTGTGTGGCATCCAGCGTGTCTCTATGGA 2436
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
434  AlaAspGlnLeuGlyTyrIleAspMetAspArgValAlaIleHisGlyTrpSerTyrGly 453
QY 2437 GGATACCTCTCCCTGATGCGCATTAATGTCAGAGGTTCAGATATCTTCAGGGTGTCTATTGCT 2496
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
454  GlyTyrLeuSerLeuMetGlyLeuValGlnTyrProGluIlePheLysValSerIleAla 473
QY 2497 GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGT 2556
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
474  GlyAlaProValThrSerTrpGluTyrTyrAspThrGlyTyrThrGluArgTyrMetAsp 493
QY 2557 CACCTTGACCAAGTCAACAGCGCTATTACTTAGGATCTGTGGCCATCGAAGCAGAAAG 2616
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
494  LeuProAspSerAsnArgSerGlyTyrAlaAlaGlySerValLeuAsnTyrIleGlnLys 513
QY 2617 TTCCCTCTGAACAAATCGTTTACTGCTCTTACATGTTTCTTCCTGGATGAGAATGTCCAT 2676
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
514  PheProAspGluAspAsnArgLeuLeuIleHisGlyLeuIleAspGluAsnValHis 533
QY 2677 TTGCGCATACACAGTATATATCTAGAGTGTATTTTAGTGGGCTGGAAAGCCATATGATTA 2736
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
534  PheHisHisThrSerGlnLeuValSerArgLeuValArgAlaAsnLysProTyrGlnLeu 553
QY 2737 CAGATCTATCTCT 2748
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
554  GlnValTyrPro 557

RESULT 14
Q54U01.DICDI PRELIMINARY; PRT; 803 AA.
ID Q54U01;
AC Q54U01;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=DD80205566;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_taxid=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachbat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nle X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churchard C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsey R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
```

RA Shauleky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.P., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RT "The genome of the social amoeba Dictyostellium discoideum.",
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL: AA01000074; EAL66689.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 803 AA; 93187 MW; 33846B370C039FAE CRC64;

Alignment Scores:
Pred. No.: 2,31e-74 Length: 803
Score: 1124.50 Matches: 291
Percent Similarity: 48.7% Conservative: 146
Best Local Similarity: 32.4% Mismatches: 310
Query Match: 20.3% Indels: 151
Dbs: 2 Gaps: 31

US-10-825-632-2 (1-3120) x Q54U01_DICDI (1-803)

QY 283 ATTGAATCAGAGATCGGCTAAATGGAGCCCTTTATGTTGAGCGGTATTCCTGGAGT 342
Db 8 IlegluGlyArgasp-----GluLeuLeuAspLysGlu 18
QY 343 CAGCTTAAAGAGCTGCTGGCGATACAGAAATATCATGGCTACATGCTGCTAAGGCA 402
Db 19 GluIleLeuLysLeuGluGlyGluGluSerGly-----SerSerAla 32
QY 403 CCACATGATTTTC-----ATGTTGTGAGAGGATGATCCA 438
Db 33 ProTyrSerPheGluPheAsnAspSerLysAsnHisLeuTyrPheLeuSerAsnValLys 52
QY 439 GATGACCTCATCTCAGACAGAAATCTATTACCTTGCCTGCTGCTGGTGAGAACAGAAAT 498
Db 53 AspGlnLysAsnIleLysAsnIleHisTyrIleAspMetAsnSerAspSerLysGluIle 72
QY 499 ACACCTGTTTATCTCGAAATCCCAAACTATCAATAGACGACGACTCTTATGCTCTCT 558
Db 73 LysProLeuPheAsnTyrValAspLysAspValGlu----- 84
QY 559 TGGAGCCTCTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAA 618
Db 85 -----LeuSerIleGlu 88
QY 619 GAAGAACTATTAGAGAAAGAAACGCAATTCGACAGCTCGGAATTCCTTACGATTAT 678
Db 89 AspGlnLeuGlnArgGluArgMetArgThrAlaAlaAsnGlyIleThrGlnPheThrPhe 108
QY 679 CACCAAGGAGTGGGAACATTTCTGTTTCAAGCCGTTAGTGGAAATTTATCACGTAAAGAT 738
Db 109 AspGlnLysHisGlnPheIleIleAlaProIleAsnAsnLysIleAsnLysIleAspIle 128
QY 739 GGAGGCCCAAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGT 798
Db 129 -----LysGluSerIleThrLysProIleLysGluIleValGlyGluThrTyrAsn 146
QY 799 CCCACATACGGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTGTTTATA 858
Db 147 HisGlnIleSerAlaAspGlyLysIle-----ValSerPheLeu 159
QY 859 CATAGCAACCATATTGGATATCTAACATCGTAACACAGAGAAGAGGAGACTCATCTAT 918
Db 160 LysAspLysAspIleThrPheIleThrAspIleSerThrAsnAlaMetTyrArgIleThrPhe 179
QY 919 GTGCAAAATGAGTACCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGTACTCTT 978
Db 180 SerAsnAspGlu-----LysHisLysPheArgTyrAlaGlyAspIleGlyPhe 196
QY 979 GTTCTCCAGAGAAATTTGATAGATATCTGGCTATTGGTGTCTCCAAAGCTGAAACA 1038
Db 197 IleTyrAlaGluGluPheSerArgTyrThrGlyTyrTrpSerProIleValGlyThr 216

QY 1039 ACTCCAGTGGTGGTAAA---ATTCTTGAATTTATATGAAGAAAATGATGAATCTGAG 1095
Db 217 CysValLysThrGlyLysProMetTyrThrIleCysTyrLeuGluGluAspGluThrAsn 236
QY 1096 GTGGAAATATTATCATGTTACATCCCTATGTTGGAAACAAGG---AGGCGAGATTCATTC 1152
Db 237 ValMetAspTyrHisIle-----ProThrSerAspLeuArgGlyLysThrThrGlnTyr 254
QY 1153 CGTTATCTCTAAACACAGGTACAGCAAT----- 1179
Db 255 LysTyrProLeuAlaGlyGluLysAsnSerIleCysLysValCysLeuValSerPheVal 274
QY 1180 -----CCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATCCTGAAGGAAGATC 1233
Db 275 LeuProThrArgThrThrPheGlnAspSerLys----- 285
QY 1234 ATAGATGTCATAGATAAGAACTAAATTCACCTTTTTCAGATTCCTATTTGAAGAGTTGAA 1293
Db 286 IleGluIleValLysSerGluLeuPheAsp---LeuLysThrGlnPheProTrpAlaGlu 304
QY 1294 TATATTGCCAGAGCTGGATGCTCTCGAGGGAATAATGCTTGGTCCATCTCTACTAGAT 1353
Db 305 TyrIleThrArgAlaGlyTrpThrProAsnGlyHisSerIleTyrLeuGlnLeuLeuAsp 324
QY 1354 CGCTCCCAAGCTCGCCTACAGATAGTGTGATCTCACTCGAATATTATTCACAGTAGAA 1413
Db 325 ArgLysGlnGlnHisLeuAlaLeuValMetValProLeuHisValPhe-----AlaGlu 342
QY 1414 GATGATGTTATGGAAGGACAGACTCATTTGAGTGCCTGATTCCTGTGAGCGCACA 1473
Db 343 AspTyrSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 360
QY 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATAATCCATGACATCTTTTCATGTTT 1533
Db 361 LeuIle---GluGluThrThrSerValTrpIleAsnIleGluPheSerPheGln---Phe 378
QY 1534 CCCCAAGTCCAGAGAGAAATTCAGTTTATTTTGCCTCTCAATGCACAAACAGGTTTC 1593
Db 379 LeuLysSerIleGluAsnGlnLeu-----IleTrpSerAsnGlu---GlnSerGlyTyr 395
QY 1594 CQTATATTATCAAAATATACATCTATTATTAAGGAAGCAAAATATAAATCCAGTCCAGTGT 1653
Db 396 ArgHisLeuTyr-----LeuIleLysTrpAspLysAsnPheThrAsnIleGln 411
QY 1654 GGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGGATAGCAATACCAGT 1713
Db 412 SerThrProIleThrLeuSer---ThrCysAsnAspAsnAspAsn-----AspAsn 427
QY 1714 GGTGAATGGAAAGTTCTTTGGCGGCATGGATCTAATATCCAGTTGATGAAGTCAGAGG 1773
Db 428 AsnAsnTrpMetVal-----SerSerAspAspIleHisIleAspGluLysArgLys 444
QY 1774 CTGCTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTACGT 1833
Db 445 LeuValTyrPheThrGlyThrLysAspThrCysLeuGluGlnHisLeuTyrValThrArg 464
QY 1834 TAGCTAAATCCT---GGAGAGGTGACAAAGGCTGACTGACCGCTGGCTACTCACAT---TCT 1887
Db 465 PheAspLysProAsnSerSerGluIleLysArgLeuSerHisAlaAsnPheSerHisArgSer 484
QY 1888 TGTGCTCAGTCAGCTGCTGACTCTTTTATAGTAGTATAGTAACTAGTAACAGAGAAGAT--- 1944
Db 485 IleSerIleSerSerAsnPheLysLysPheIleThrThrTyrSerAsnIleSerThrIle 504
QY 1944 ----- 1944
Db 505 SerLysThrGluValPheAspLeuIleTyrAsnAspAspAsnAsnAspAsnAspIle 524
QY 1945 ---CCACACTGTGTGCTCTTTTACAGCTATCAAGTCTGAAGATGACCCCACTTGCAAA 2001
Db 525 TyrProIleValLysSerSerPhePheIleAsnAspAspAspAspAspAspAspAspAsp 544

Qy	2002	ACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTCTCTGACTATACCTCT	2061	RA	Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,		
Db	545	LysLysLysIleAsnIleAsnIle-----	552	RA	Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,		
Qy	2062	CCAGAAATTTTCTCTTTTAAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAG	2121	RA	Attix C., Andreise T., Frankheim M., Amico-Keller G., Coefield J.,		
Db	553	ProLysIlePheAsnPhelysAsnSerlysglyValThrIleTyrlleGlyGlnTyThrLeu	572	RA	Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,		
Qy	2122	CCTCATGATCTACAGCCTGGAAGAAATATCTACTCTGCTGCTTCATATATGCTGCTCT	2181	RA	Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,		
Db	573	ProSerAspTySerlylsAsplyslystyrProThrValValTyThrValTyThrGlyPro	592	RA	Olsen A.S., Carrano A.V.,		
Qy	2182	CAGTGCAGTTCGTGATATCGGTTTAAAGAGTCAGTATTTCCGCTTGAATACCCCTA	2241	RL	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.		
Db	593	HisValGlnIleValArgAsnGlnTyraenTyrllelysglnHisTyThrAsn-----	610	DR	EMBL; AC005783; AAC62840.1; -; Genomic DNA.		
Qy	2242	GCCTCTCTAGTTCGTGCTGATGATAGACACAGGGGATCTCTCACCAGGGCTT	2301	DR	Ensembl; ENSG00000142002; Homo sapiens.		
Db	611	-----PheGlyPheIleGlnValMetIleAspAsnValGlySerAlaAsnArgGlyLeu	628	DR	InterPro; IPR002469; Peptidase_S9B.		
Qy	2302	AAATTTGAAGCGCCCTTTAAATATAAATGGGTCAAAATGAGAAATTCAGCATCAGGTGGA	2361	DR	Pfam; PF00930; DPPIV_N; 1.		
Db	629	GlupheGlySerHisIleArgGlnlyMetGlyGlnValGlnIleGlyAspGlnValGlu	648	FT	NON_TER 432 432		
Qy	2362	GGACTCCAATATCTAGTTCCTCGATATGATTTCT-----ATTGACTTAGATCGTGGGC	2415	SQ	SEQUENCE 432 AA; 48595 MW; 64E2B85BE0523A7E CRC64;		
Db	649	GlyIleAsnTyrlleuValGlyAsn---AspIleValSerIleAspValAsnArgIleAla	667	Alignment Scores:			
Qy	2416	ATCCAGCGTGTCTATGAGGATACCTCTCCCTGATGGCAATTAATGCGAGGTGAT	2475	Pred. No.:	7,76e-69 Length: 432		
Db	668	IleSerGlyTrpSerTyrglyTyraenSerLeuMetAlaIleSerGlnArgProAsp	687	Score:	1049.00 Matches: 189		
Qy	2476	ATCTCAGGTTGCTATGCTGGGCCAGTCATCTCTGCTGATCTTCTATGATACAGGA	2535	Percent Similarity:	74.3% Conservative: 63		
Db	688	ValPheLysIleAlaValCysGlyAlaProValSerAspTrpArgLeuTyraenThrGly	707	Best Local Similarity:	55.8% Mismatches: 79		
Qy	2536	TACACGGAACGTTTATGCGTCACTGACAGATGACAGGCTTACTTAGGATCT	2595	Query Match:	18.9% Indels: 8		
Db	708	TyrThrGluArgTyMetAsnValProGlnAspAsnIleAspGlyTyrlleLysLeuGlyAsp	727	DB:	2		
Qy	2596	GTGGCCATGCAACAGAAAGTTCCTCTGAAACAAATCTGTTACTGCTCTTACATGGT	2655	US-10-825-632-2 (1-3120) x 075868_HUMAN (1-432)			
Db	728	ThrThrHisTySer-----PheProThrGluGluAsnArgLeuLeuIleHisGly	745	Qy	316	TTTATGTTGAGCGGTATTCTTGAGTTCAGCTTAAAGCTGCTTCCGATACACAGAAA	375
Qy	2656	TTCTGTGATGAGATCTCCATTTTGACATACCATGATATATCTAGTATTTTAGTAGG	2715	Db	95	PheGlnValGlnlyHisSerTrpAspGlyLeuArgSerIleIleHisGlySerArglys	114
Db	746	LeuGlnAspGluAsnValHisPheSerAsnThrIleTyrlleIleAspHisLeuThrLys	765	Qy	376	TATCATGGCTACATGATGCTAAGCCACCATGATTTTCATGTTTGTGAAGAGGAATCAT	435
Qy	2716	GCTGGAAGCCATGATTTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCT	2775	Db	115	TyrSerGlyLeuIleValAsnlyAspPheGlnPheValGlnlyThrAsp	134
Db	766	ThrGlnlySerProTyrlleLeuLyThrLeuProAsnGluArgHisGlyValArgAsnThr	785	Qy	436	CCAGATGGAGCTCATTCAGACAGATCTATTACCTTGCCATGCTCTGTGGTGAACAGAGAA	495
Qy	2776	GAATCGGAGAACATTATGACATGTCATCTTTTGCATCTACTCTTCAAGAAACCTT	2829	Db	135	GluSerGlyProHisSerHisArgLeuTyrlleGlyMetProTyrlleSerArgGlu	154
Db	786	AspAsnArgIleTyrlleGlyLeuPheValIleAsnHisLeuLeuLysAsnLeu	803	Qy	496	AATACACTGTTTATTCTGAAATTCCTCAAACTATCAATAGACAGCAGTCTTAAATGCTC	555
RESULT 15				Db	155	AsnSerLeuLeuTyrlleSerGluIleProlyLysValArgLysGluAlaLeuLeuLeu	174
O75868 HUMAN	ID	O75868 HUMAN PRELIMINARY; PRT; 432 AA.		Qy	556	TCITGGAAGCTCTTTTGGATCTTTTTCAGGCAACACATGGACTATGGAATGATTTCTCGA	615
AC	O75868			Db	175	SerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisHisGlyValTySerArg	194
DT	01-NOV-1998 (TrEMBLrel. 08, Created)			Qy	616	GAAGAAGACTATTAAAGAGAAAGAACCCATTTGGAACAGTCGCGAATGCTCTTACGAT	675
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)			Db	195	GluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIleThrSerTyAsp	214
DT	01-NOV-1998 (TrEMBLrel. 08, Last annotation update)			Qy	676	TATCACCAGAGAGTGGACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAA	735
DE	R33083.1 (Fragment)			Db	215	PheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheHisCysArg	234
OS	Homo sapiens (Human)			Qy	736	GATGAGGGGCCCAAGGATTT-----ACGCAACACCTTTTAAGG	774
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			Db	235	AspGlyGlyLysAsnGlyPheMetValSerProGlyProGlyCysValSerProMetLys	254
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			Qy	775	CCCAATCTAGTGGAACTAGTTCTGCCAACATACGATGATCCAAATATTCGCCCGCT	834
OC	Homo.			Db	255	ProLeuGluIleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAla	274
OX	NCBI_TaxID=9606;			Qy	835	GATCAGACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAAC	894
RN	(1)			Db	275	AspProAlaPhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThr	294
RA	Lamerdin J.E., McCreedy P.M., Skowronski E., Viewanathan V.,			Qy	895	AGAGAAGAAAGGAGACTCATCTTATGTCACATAGCTAGCCAAACATGAGAGAGATGCC	954
RA	Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,			Db	295	GlyGluGluArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspPro	314
RP	NUCLEOTIDE SEQUENCE.			Qy	955	AGATCAGCTGAGTCGCTACCTTTGTTTCCCAAGAGAAATTTGATAGATATTTCTGGCTAT	1014
RA				Db	315	LysSerAlaGlyValAlaThrPheValIleGlnGluGluPheAspArgPheThrGlyTyr	334
RA				Qy	1015	TGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAAATTTCTTAGAATTTCTATAT	1074

Db				:::						353
	Trp	Trp	Cys	Pro	Thr	Ala	--	Ser	Trp	Glu
	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu
Qy	1075	GA	CA	AA	AT	GAT	GA	TCT	GAG	TGG
		AA	AT	GA	TCT	GAG	TGG	AA	AT	TAT
		TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT
		TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT
Db	354	Glu	Glu	Val	Ala	Pro	Ser	Pro	Ala	Leu
		Glu	Glu	Val	Ala	Pro	Ser	Pro	Ala	Leu
		Glu	Glu	Val	Ala	Pro	Ser	Pro	Ala	Leu
Qy	1135	AG	GA	GG	CG	AG	AT	TCT	CT	TA
		AG	GA	GG	CG	AG	AT	TCT	CT	TA
		AG	GA	GG	CG	AG	AT	TCT	CT	TA
		AG	GA	GG	CG	AG	AT	TCT	CT	TA
Db	374	Arg	Gly	Thr	Ala	Pro	Ser	Pro	Ala	Leu
		Arg	Gly	Thr	Ala	Pro	Ser	Pro	Ala	Leu
		Arg	Gly	Thr	Ala	Pro	Ser	Pro	Ala	Leu
Qy	1195	AA	GA	TG	CA	GA	AA	TA	AT	GAT
		AA	GA	TG	CA	GA	AA	TA	AT	GAT
		AA	GA	TG	CA	GA	AA	TA	AT	GAT
		AA	GA	TG	CA	GA	AA	TA	AT	GAT
Db	394	Lys	Leu	Ala	Glu	phe	Gln	Thr	Ala	Glu
		Lys	Leu	Ala	Glu	phe	Gln	Thr	Ala	Glu
		Lys	Leu	Ala	Glu	phe	Gln	Thr	Ala	Glu
Qy	1255	CT	AA	TT	CA	CC	TT	TT	GA	CT
		CT	AA	TT	CA	CC	TT	TT	GA	CT
		CT	AA	TT	CA	CC	TT	TT	GA	CT
		CT	AA	TT	CA	CC	TT	TT	GA	CT
Db	414	Leu	Val	Gln	Pro	Phe	Ser	Leu	Phe	Pro
		Leu	Val	Gln	Pro	Phe	Ser	Leu	Phe	Pro
		Leu	Val	Gln	Pro	Phe	Ser	Leu	Phe	Pro

Search completed: April 14, 2006, 11:31:04
 Job time : 779 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: April 14, 2006, 11:31:23 ; Search time 18.6 Seconds
(without alignments)
2773.636 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5552
Sequence: 1 aagtcgaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Listing first 45 summaries

Post-processing: Minimum Match 0%

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlpl
-O=/abs/ABSSWEB_spool/US10825632/runat_14042006_092353_3529/app_query.fasta_1
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10825632 @CNG 1.1 71 @runat_14042006_092353_3529 -NCPU=6
-ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/FCUTS COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	84.7	882	US-09-976-674-1	Sequence 1, Appli
2	4700	84.7	882	US-10-070-464-1	Sequence 1, Appli
3	3607.5	65.0	690	US-09-976-674-7	Sequence 7, Appli
4	3513.5	63.3	661	US-09-976-674-11	Sequence 11, Appl
5	3504	63.1	658	US-09-976-674-19	Sequence 19, Appl
6	3236	58.3	613	US-09-976-674-21	Sequence 21, Appl
7	2871	51.7	892	US-09-976-674-23	Sequence 23, Appl
8	2871	51.7	892	US-09-976-674-27	Sequence 27, Appl
9	2870	51.7	863	US-09-976-674-3	Sequence 3, Appli
10	2821.5	50.8	879	US-09-976-674-33	Sequence 33, Appl
11	2821.5	50.8	879	US-09-976-674-35	Sequence 35, Appl
12	2422	43.6	465	US-10-070-464-5	Sequence 5, Appli

13	2407	43.4	832	2	US-09-976-674-29	Sequence 29, Appl
14	2407	43.4	832	2	US-09-976-674-31	Sequence 31, Appl
15	2357.5	42.5	819	2	US-09-976-674-37	Sequence 37, Appl
16	2357.5	42.5	819	2	US-09-976-674-39	Sequence 39, Appl
17	1836.5	33.1	360	2	US-10-070-464-7	Sequence 7, Appli
18	1808	32.6	358	2	US-09-976-674-13	Sequence 13, Appl
19	1645.5	29.6	310	2	US-09-794-236-4	Sequence 4, Appli
20	1645.5	29.6	310	2	US-10-070-464-3	Sequence 3, Appli
21	1392	25.1	518	2	US-09-976-674-25	Sequence 25, Appl
22	1278	23.0	241	2	US-09-976-674-9	Sequence 9, Appli
23	1007.5	18.1	194	2	US-09-976-674-17	Sequence 17, Appl
24	724.5	13.0	981	2	US-09-902-540-16812	Sequence 16812, A
25	616	11.1	710	2	US-09-518-550-28	Sequence 28, Appl
26	582.5	10.5	771	2	US-09-462-284-2	Sequence 2, Appli
27	582.5	10.5	771	2	US-09-079-592-2	Sequence 2, Appli
28	572.5	10.3	723	2	US-09-518-550-29	Sequence 29, Appl
29	572	10.3	732	2	US-09-518-550-30	Sequence 30, Appl
30	569.5	10.3	676	2	US-09-518-550-42	Sequence 42, Appl
31	529	9.5	766	2	US-10-002-593-6	Sequence 6, Appli
32	529	9.5	766	2	US-09-949-016-6146	Sequence 6146, Ap
33	529	9.5	766	2	US-10-423-714-6	Sequence 6, Appli
34	525	9.5	755	4	PCT-US93-07923-3	Sequence 3, Appli
35	525	9.5	759	4	PCT-US93-07923-2	Sequence 2, Appli
36	525	9.5	766	1	US-08-230-491A-3	Sequence 3, Appli
37	525	9.5	766	1	US-08-619-280A-3	Sequence 3, Appli
38	525	9.5	766	1	US-08-940-391-3	Sequence 3, Appli
39	525	9.5	766	2	US-09-794-236-1	Sequence 1, Appli
40	525	9.5	766	2	US-09-265-606-3	Sequence 3, Appli
41	525	9.5	775	2	US-09-949-016-10450	Sequence 10450, A
42	519	9.3	766	2	US-09-518-550-27	Sequence 27, Appl
43	488.5	8.8	737	2	US-09-902-540-11421	Sequence 11421, A
44	484.5	8.7	818	2	US-09-462-845-3	Sequence 3, Appli
45	484.5	8.7	818	2	US-10-402-312-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Alignment Scores:
Pred. No.: 0
Score: 4700.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 84.7%
DB: 2
Length: 882
Matches: 882
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0
US-10-825-632-2 (1-3120) x US-09-976-674-1 (1-882)

QY 214 ATGCGACGACAAATGGAAACAGACAGCTGGGTGTGAGATATTTTGAACCTGCGACGTGT 273
|||||
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

QY	274	GAGGAGAAATATTGAAATCACAGGATCGGCCTAAATTGGAGCCCTTTTATATGTTGAGCGGTAT	333	QY	1354	CGCTCCAGACTCGCCTACAGATAGTGTGTATCTCACCTGAATTTATTTATCCAGTAGAA	1413
DB	21	GluGluAsnIleGluSerGlnAspA-gProLysLeuGluProPheTyValGluArgTyr	40	DB	381	ArgSerGlnThrArgLeuGlnIleValIleSerProGluLeuPheIleProValGlu	400
QY	334	TCCTGGAGTCAGCTATAAAAGCTGCTTGGCGATACCCAGAAATATATCATGGCTACATGATG	393	QY	1414	GATGATCTTATGGAAGCGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGAGCGCCACTA	1473
DB	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60	DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	394	GCTAAGGACCACATGATTTTCATGCTTTGTGAAGAGGAATGATCCAGATGGACCTCATPCA	453	QY	1474	ATTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT	1533
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80	DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	454	GACAGAATCTATTACCTTGGCATGCTCTGGTGAGAACAGAGAAAAATACACTGTTTTATTCT	513	QY	1534	CCCCAAGTCACGAAGAGAAATTGAGTTATTTTTCCTCTGCTGAATCCAAAAACAGGTTTC	1593
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100	DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	514	GAATATCCCAAACTATCAATAGACGACGAGCTTAAATGCTCTCTTGGAGGCTCTTTTG	573	QY	1594	CCTCATTTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAACCATCCAGTGGT	1653
DB	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120	DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	574	GATCTTTTCAGGCAACACTGGATATGGAATGATATCTCGAGAGAGAACTATTAGA	633	QY	1654	GGCTCGCTCGCTCCCAAGTATTCAAGTGTCTATCAAGAGGAGATAGCAATATCCAGT	1713
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140	DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	634	GAAGAAACCCATGGACAGTCGGAATTTTATCAGTAAAGATGGAGGGCCACACAGA	753	QY	1714	GTGTAATGGAAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAGG	1773
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	694	ACATTTCTGTTCGAAGCGGAGTGGAAATTTTATCAGTAAAGATGGAGGGCCACACAGA	753	QY	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
QY	754	TTTAGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG	813	QY	1834	TACGTAAATCTCGGAGAGGTGACAGGCTGACTCACCTGGCTACTCACATTTCTTGTCTGC	1893
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	814	GATCCAAATATTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACCATATT	873	QY	1894	ATCAGTCAGCATGTGACTCTTTATTAAGTAAGTATAGTAACACAGAGAATCCACACTGT	1953
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	DB	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	874	TGGATATCTAATCGTAACACAGAGAAGAAGAGACTCACTTATGTGCACAAATAGACTA	933	QY	1954	GTGTCCTTTTACAAGCTATCAAGTCTCAAGATGACCCAACTCTGCAAAAAACAAGAAATTT	2013
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	DB	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	934	GCCACATGGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTTCTCCAAGAGAA	993	QY	2014	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATACTCTCCAGAAATTTTC	2073
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
QY	994	TTTGATAGATATCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACCTCCAGTGGTGT	1053	QY	2074	TCCTTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	2133
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	DB	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	1054	AAAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATTCATGTT	1113	QY	2134	CAGCTGGAAAGAAATATCTCTACTCTGTGTTTCATATATGTTGTTCTCAGGTGCAGTTG	2193
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	DB	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	1114	ACATCCCTTATGTTGGAAACAGGAGGCGAGATTTCATTCGTTATCTTAAAAACAGGTACA	1173	QY	2194	GTGAATANTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTTAGGT	2253
DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	DB	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	1174	GCAATCTCTAAAGTCACTTTTAAAGATGTCAAGAAATATGATGATGCTCAAGGAGGATC	1233	QY	2254	TATGTGGTGTGATGATAGACAAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTGAAGGC	2313
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	DB	681	TyrValValValIleAspAsnArgGlySerCysHisArgLysLeuLysPheGluGly	700
QY	1234	ATAGATGTCATAGATAAGGAATAATTCAACCTTTTGAGATTTCTATTGAGGAGTTGAA	1293	QY	2314	GCCTTTAAATATATAAATGGGTCAATAGAAATGACATGACAGGTGGGAAGGATCTCAATAT	2373
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	DB	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	1294	TATATTGCCAGAGCTGGATGACTCTCGAGGAGAAATATCTTGGTCCATCTCTACTAGAT	1353	QY	2374	CTAGCTTCTCGATATATGATTTTCACTTAGATCTGTGGGATCCACCGGTGGTCTTAT	2433
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	DB	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
				QY	2434	GGAGGATACCTCTCCCTGATGGCATTTAATGCAGAGGTGAGATATCTTCCAGGTGCTTAT	2493

Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Qy 2494 GCTGGGCCCCAGTCACCTCTGTGGATCTTATGATACAGGATACACGGAAGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Qy 2554 GGTCACTCCAGAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlu 800
Qy 2614 AAGTTCCCTCTGAACCAATCGTTTACTCTCTTACATGTTTCTCGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Qy 2674 CATTTTGCATACACAGTATATTACTAGTGTGTTTATAGTGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Qy 2734 TTACAGATCTATCTCAGGAGACACAGCATTAAGAGTTCTCGAATCGGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Qy 2794 GAACGTGATCTTTTGCATCTACCTTCAAGAAAACTTGGATCAGTATTCCTGCTCTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Qy 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 2

US-10-070-464-1
; Sequence 1, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-1

Alignment Scores:

Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-1 (1-882)

Qy 214 ATGGCAGCAGCAATCGGAACAGACACCTGGGTGTCAGATATTTGAACTCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAGAATATTGAATCACAGGATCGGCCTAAATTTGGAGCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40

Qy 334 TCCTGGAGTCAGCTTTAAAAAGCTGCTCCGATACCAGAAATATCATGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGGCCACCATGATTTCAATGTTGTGAAGAGGAATGATCCAGATGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy 454 GACAGAAATCTATTACCTTCCCATGCTCGTGAAGAACAGAGAAATACACATGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy 514 GAAATTTCCCAAAATCATCAATAGACAGCAGTCTTAATGCTCTCTTGAAGCCCTTTTGT 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy 574 GATCTTTTTCAGCAACACTGAGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Qy 634 GAAAGAAAAACGCAATTCGAACAGTCGGAATTCCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Qy 694 ACATTTCTGTTTCAAGCCGTTAGTGAATTTATCACGTAAAGATGAGAGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Qy 754 TTTAGCCAAACCTTTTAAGGCCCAATCTAGTGGAACTAGTTTCTCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Qy 814 GATCCAAATATTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Qy 874 TGGATATCTAATCATCGTAACAGAGAAGAAAGAGACTCCTATGTATGTCACAATGAGTA 933
Db 221 TrpIleSerAsnIleValThrArgGluArgLeuThrTyrValHisAsnGluLeu 240
Qy 934 GCCAATGGAAGAAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTTCTCAAGAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Qy 994 TTTGATAGATATTCGGCTATTGGTGTCTCAAAAGCTGAAACAACTCCAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Qy 1054 AAAATTTCTAGAAATCTATATGAAGAAATGATGATCTGAGGTGGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Qy 1114 ACATCCCTTATGTTGGAACCAAGGAGGCCAGATTCATTCCCTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Qy 1174 GCNAATCTTAAAGTCATTTTAAGATGTCAGAAATATGATGATGCTGAGGAAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Qy 1234 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCATTGTAAGGAGTGA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy 1294 TATATTGCCAGAGCTGGATGGACTCCTCGAGGAAAATATGCTTGGTCCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Qy 1354 CGCTCCAGACTCGCCTACAGATGTTGATCTCACCTGTAATTTATTTATCCCAAGTGA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400

QY	1414	GATGATGTTATGAAAGGAGGAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA	1473	Db	761	AlaGlyAlaProValThrLeuThrPhePheArgThrGlyThrGluArgTyrMet	780
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	QY	2554	GGTCACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTCTGCCATGCAAGCAGAA	2613
QY	1474	ATTATCTATCAAGAAACAACAGACATCTCGGATAAATATCCATGACATCTTTTCATGTTTTT	1533	Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	421	IleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe	440	QY	2614	AAGTTCCCTCTGAAACCAAAATCGTTTACTCTCTTACATGGTTTCTCGATGGAATGTC	2673
QY	1534	CCCCAAGTCACCAAGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC	1593	Db	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	QY	2674	CATTTTGCACATACCAGTATATTAAGTATTTTGTAGTGGGCTGGAAGCCATATGAT	2733
QY	1594	CGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1653	Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	QY	2734	TTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCCTGAATCGGAGAACATTAT	2793
QY	1654	GGGTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGGATAGCAATACCAAGT	1713	Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	QY	2794	GAACTGCATCTTTTGCATCTACCTTCAAGAAAAACCTTGGATCAGGTATTGCTCTAAAA	2853
QY	1714	GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCAGAAG	1773	Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	QY	2854	GTGATA 2859	
QY	1774	CTGCTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT	1833	Db	881	ValIle 882	
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540	RESULT 3			
QY	1834	TACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGTACTCACATTTCTGTGTC	1893	US-09-976-674-7			
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	; Sequence 7, Application US/09976674			
QY	1894	ATCAGTCAGACTGCTGACTCTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953	; Patent No. 6844180			
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	; GENERAL INFORMATION:			
QY	1954	GTGTCCTTTACAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAAGAAATTT	2013	; APPLICANT: Q1, Steve			
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	; APPLICANT: Akinsanya, Karen			
QY	2014	TGGGCCACATTTTGGATTCAGCAGGTCTCTCTCTGACTATACTCTCCACAAATTTTC	2073	; APPLICANT: Riviere, Pierre			
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	; APPLICANT: Junien, Jean-Louis			
QY	2074	TCTTTTGAAGTACTACTGGATTTTACATGTTATGGGATGCTCTACAAGCCTCATCTATA	2133	; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	; FILE REFERENCE: 70669			
QY	2134	CAGCTGGAAGAAATATCTCTGCTGCTGTTTATATATGCTGCTCTCAGTGCAGTGTG	2193	; CURRENT APPLICATION NUMBER: US/09/976,674			
Db	641	GlnProGlyLysLysTyrProThrValIlePheIleTyrGlyGlyProGlnValGlnLeu	660	; PRIOR FILING DATE: 2001-10-12			
QY	2194	GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGTTGAATACCCCTAGCCTCTAGGT	2253	; PRIOR APPLICATION NUMBER: US 60/240,117			
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680	; PRIOR FILING DATE: 2000-10-12			
QY	2254	TATGTGGTTGTAGTGATAGACAACAGGGGATCTGTCCACCGAGGCTTAAATTTGAAGC	2313	; NUMBER OF SEQ ID NOS: 61			
Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700	; SOFTWARE: PatentIn version 3.1			
QY	2314	GCCTTTAAATATAAAATGGGTCAATAGAAATTCAGCATCAGGTGGAAGACTTCCAATAT	2373	; SEQ ID NO 7			
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720	; LENGTH: 690			
QY	2374	CTAGCTTCTCGATATGATTTTCATTTAGCTTAGATCGTGTGGCATCCACGGCTGGTCTAT	2433	; TYPE: PRT			
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyr	740	; ORGANISM: Homo sapiens			
QY	2434	GGAGGATACCTCTCCCTGATGCAATTAATGCAAGGTTCAGATATCTTCAGGTTGCTATT	2493	US-09-976-674-7			
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760	Alignment Scores:			
QY	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATG	2553	Pred. No.: 0			
				Score: 3607.50			
				Percent Similarity: 87.1%			
				Best Local Similarity: 87.1%			
				Query Match: 65.0%			
				DB: 2			
				US-10-825-632-2 (1-3120) x US-09-976-674-7 (1-690)			
				QY 214 ATGGCAGCAGCAATGGAACACAGACAGCTGGTGTGAGATATTTGAAACTGCGACTGT			
				Db 1 MetAlaAlaMetGluThrGlnGluLeuGlyValGluIlePheGluThrAlaAspCys			
				QY 274 GAGGAGAATTTGAATACAGGATCGGCTCTAAATTTGGAGCCCTTTTATTTGAGGGGTAT			
				Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr			
				QY 334 TCCTGGAGTCAGCTTAAAAAGCTGTGTCGCGATACACAGAAATATCATGGCTACATGATG			
				Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet			
				QY 394 GCTAAGGCACCATCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGAGCTCATTC			

Db 61 AlaIysAlaProHisAspPheMetPheValIysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACTTGGCCATGTCTGGTGAGAACAGAGAAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAACTATCAATAGACGACAGTCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGATTTGGAACTCGGAATTCCTTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTCTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGGCCACAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCGCAACACCTTTAAGCCCMATCTAGTGAAACTAGTGTGCCAATACCGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATGATCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TCGATATCTAACATCGTAACAGAGAGAAAGGAGACTCACTTATGTGCACATAGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGlnGluArgGluLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATCGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATPAGATATTCTGGCTATTGTGGTGTCCAAAGCTGAAACACTCCCACTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTGAATTTCTATAGAAGAAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACAAGAGGGCCAGATTCAATCCGTTATCCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAATCTTAAGTACATTTTAAGATGTCAGAAATATGATGATGCTGAAGGAAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCCTATTGAGAGGTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGACTGGATGGACTCCTCAGGGAATAATGCTTGGTCCATCCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGTCCCGACACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAGAACAACACAGACATCTGGATAATATCCATGACTTTTCATGTTTTT 1533

Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCACGAAGAGCAAAATTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTATACAAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAATACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAGAGGAGATGACGAATACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTCGCCGCCATGAGTCTATATCAAGTTGATCAAGTCAGAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCCTGTAGTACTAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACGTAAATCCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACTTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAGAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCAATTTGGATTCAGCAGGTCTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATATCTACTGTGCTGTTTCAATATATGTTGGTCTCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
QY 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGCTTGAATACCTCTAGCTCTCTAGGT 2253
Db 655 655
QY 2254 TATGTGTTGTAGTGATAGACAACAGGGGATCCTGTACCGAGGGGCTTAAATTTGAAGGC 2313
Db 655 655
QY 2314 GCCTTTAAATATAAATGGGTCAATAGAAAATTGACGATCAGGTGGAAGGACTCCAATAT 2373
Db 655 655
QY 2374 CTAGCTTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTAT 2433
Db 655 655
QY 2434 GGAGGATACCTCTCCCTGTATGGCATTAATGTCAGAGGTGAGATATCTTCAGGGTGTGCTATT 2493
Db 656 659
QY 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db 659 uLeuGlyProGlnSerLeuCysGlySerSerMetIleGlnAspThrArgAsnValIleTr 679
QY 2554 GGTCAACCTGACCAAGATGAACAGGCGCTATTACT 2587
Db 679 pValThrLeuThrArgMetAsnArgAlaIleThr 690

Tue Apr 18 08:18:35 2006

```
RESULT 4
US-09-976-674-11
; Sequence 11, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akingsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-11

Alignment Scores:
Pred. No.: 0 Length: 661
Score: 3513.50 Matches: 660
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 63.3% Indels: 2
DB: 2 Gaps: 1

US-10-825-632-2 (1-3120) x US-09-976-674-11 (1-661)
QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGGTGGAGATATTTGAAACTCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluLeuGluValGluLeuPheGluThrAlaAspCys 20
QY 274 GAGGAGATATTTGATCAGAGGATCGGCTAAATGGAGCCTTTTATCTTCGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProIysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTCTTCCGATACCGAATAATATCATCGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCCACCATGATTTTCATGTTGTGAAGGAGATGATCCAGATGGACTCATTTCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTCTATTACCTTCCCATGCTCTGGTGAGAACAGAGAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCACAAACTATCAATAGACGACAGCTCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTCGACTATTGGAATGATTCTCGAGAAGAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAACGATTTGGACAGTCGGAATTCCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgGileGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTTCAGCGGTAGTGGATTTTATCACCTTAAAGATGAGGCGGCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTAAGCCCAATCTAGTGAACACTAGTGTGTCACACATACCGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
814 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATATACATACGACGATATT 873
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
874 TGGATATCTAAACATCGTTAACCCAGAGAAGAGAGACTCACTTATGTGCAATGAGCTA 933
221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
934 GCCAACATCGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTGTCTCCAAAGAA 993
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
994 TTTTGATAGATATTTCTGGCTATTGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGT 1053
261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
1054 AAAAACTTTAGAAATCTATATGAAGAAAAATGATCAATCTCAGGTGGAAAAATTTATCATGTT 1113
281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
1114 ACATGCCCTATGTTGGAAACAGAGGCGCAGATTTCATTCGCTTATCTTAAACAGGTACA 1173
301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
1174 GCAATCTTAAAGTCACCTTTTAAAGTGTCAAGAAATATGATGATGCTGCTGAAGGAAGATC 1233
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTTCAGATCTCTTATTTGAAGAGTTGAA 1293
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
1294 TATATTGCCAGAGCTGGATGAGTCTCTGAGGAAAAATATGCTGTGCTCCATCTACTAGAT 1353
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
1354 CGCTCCCAACACTCGCTCAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1413
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
1414 GATGATGTTATCGAAAGGAGCAGAGACTCATTCAGTCAGTCCTGATTTCTGTGACCCCACTA 1473
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
1474 ATTATCTGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTT 1533
421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
1534 CCCAAAGTCAGAGAGAGGAAATTCAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTC 1593
441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATTAACGATCCAGTGGT 1653
461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
1654 GGGCTGCTGCTCCAAAGTGAATTCAGTGTCTTATCAAGAGAGAGATGACAAATACCAGT 1713
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleHisSer 500
1714 GGTGAATGGGAAGTCTTTCGCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAAGG 1773
501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
1774 CTGATATATTTTGAAGGCGCAAAAGACTCCCTTTTAGAGCATCACCTGCTACGTAGTCA 1833
521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
1834 TACGTAATCTCGAGAGGTGCAAGCGTGTGACCGCTGCTGCTACTCACTTCTTGTGTC 1893
541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
```

1894 ATCAGTCAGCAGTGTGACTTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db |||||
561 IleserGlnHisCysaspPhePheSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy |||||
1954 GTGTCCCTTTTACCAAGCTATCAAGTCTCTGAAGATGACCCCAACTGTGCMAAACAAAGAAATTT 2013
Db |||||
581 ValSerLeuTyrLysLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Qy |||||
2014 TGGGCCACCAATTTTGGATTACAGAGTCTCTCTCTCTGCTATATCTCTCCAGAAATTTTC 2073
Db |||||
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
Qy |||||
2074 TCTTTTCAAGTACTACTGTGATTACATTTGATGGAGTCTCTACCAAGCCTCATGACTA 2133
Db |||||
621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Qy |||||
2134 CAGCTGGAAGAAATATCTACTGCTGCTTCATATATGGTGGTCTC--AGTGGCAG 2190
Db |||||
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly-LeuLeuArgCysSe 660
Qy |||||
2191 TTGG 2194
Db |||||
660 ITrp 661
RESULT 5
; Sequence 19, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19
Alignment Scores:
Pred. No.: 0 Length: 658
Score: 3504.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.1% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x US-09-976-674-19 (1-658)
Qy 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTGGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAGATATTGAATCACAGATCGGCTAAATTTGAGCCTTTTATGTGTAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Qy 334 TCCTGGAGTCAGCTTTAAAGCTGCTTGGCGATACCAAGAAATATCATGCTCATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGGCCACCATGATTTTATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 453
Db |||||
61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

Qy 454 GACAGAATCTATTACCTTGGCCATGCTGTGGTGAGACAGAGAAAATACACTGTTTATTCT 513
Db |||||
81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy |||||
514 GAAATTTCCAAAACCTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Db |||||
101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy |||||
574 GATCTTTTTCAGGCAACACTGCACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAGA 633
Db |||||
121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Qy |||||
634 GAAAGAAAACGATTTGGAACAGTCCGAAATTCCTTTACGATTATACCAAGAAAGTGGGA 693
Db |||||
141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Qy |||||
694 ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCACGTAAAGATGAGAGGGCCACAGGA 753
Db |||||
161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Qy |||||
754 TTTAGCGAAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCACATACGATG 813
Db |||||
181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Qy |||||
814 GATCCAAAATATATGCCCGCTGATCCAGACTGGATTGCTTTATACATAGCAACGATATT 873
Db |||||
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Qy |||||
874 TGGATATCTAACATCGTTAAACAGAGAAAGAGAGACTCACTTATGTGCACAATAGCTA 933
Db |||||
221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Qy |||||
934 GCCAATGGAAGAAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAAAGAGAA 993
Db |||||
241 AlaAsnMetGluGluAspAlaAspSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Qy |||||
994 TTTGATAGATATCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGT 1053
Db |||||
261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Qy |||||
1054 AAAATTTTAGAATCTATATGAAGAAATGATGAACTGAGGTGGAAATTTATTCATGTT 1113
Db |||||
281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Qy |||||
1114 ACATCCCTTATGTTGAAACAAGGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACA 1173
Db |||||
301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Qy |||||
1174 GCAATCTCTAAAGTCACTTTTAAGATGTGAGAAATATGATTTGATGTGAAGAGGATC 1233
Db |||||
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Qy |||||
1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGTAGATTCCTTTTGAAGAGGTGAA 1293
Db |||||
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy |||||
1294 TATATTGCCAGACTGGATGGAATCTCTGAGGAAAATATGCTTGTGTCATCTACTAGAT 1353
Db |||||
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Qy |||||
1354 CGTCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Db |||||
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Qy |||||
1414 GATGATGTTATGAAAGGAGCAGACTCATTTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
Db |||||
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Qy |||||
1474 ATTATCTATCAAGAAAACACAGACATCTCGATATAATATCCATGACATCTTTTCATGTTT 1533
Db |||||
421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Qy |||||
1534 CCCCAAGTCACCAAGAGGAAATTGAGTTTATTTTTCCTCTGAAATGCAAAACAGGTTC 1593

Db	441	ProGlnSerHisGluGluGluGluGluPhePheAlaSerGluCysLysThrGlyPhe	460	US-10-825-632-2 (1-3120) x US-09-976-674-21 (1-613)	
Qy	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGTCAGTGGT	1653	214	ATGGCAGCAGCAATGGAAACAGACACACTGGTCTTCAGATATTTCGAACCTGGGACTGT 273
Db	461	ArgHisLeuTyrLysIleThrSerLysLysLysGluSerLysLysArgSerSerGly	480	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy	1654	GGGCTGCTCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATACCGT	1713	274	GAGGAGATATTGAATTCACAGATCGGCTTAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Qy	1714	GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCAGAAGG	1773	334	TCTCTGGAGTCACTTAAAGACCTGCTTCGCGATACCAAGAAATATCATGCTCATGATG 393
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy	1774	CTGTATATTTTGAAGCACCAAGACTCCCTTTAGACATCACTGTACGTAGTCACT	1833	394	GCTAAGGCCACCACTGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATPCA 453
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSer	540	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy	1834	TACGTAATCTCGAGAGGTGACAAAGCTGACCTGACCGGTGCTACTCATCTTGTCTGC	1893	454	GACAGAATCTATTACCTTGGCCATGCTGTGGTGAGAACAGAGAAATACACTGTTTATTCT 513
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy	1894	ATCAGTCAGCACTGTGACTTCTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGT	1953	514	GAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG 573
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy	1954	GTGTCCCTTTTACAAAGCTATCAAGTCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTT	2013	574	GATCTTTTTCAGGCAACACTCGACTATGGAATGTATTCTCGAAGAAGAACTATTAAAGA 633
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Qy	2014	TGGGCCACCACTTTTGGATTTCAGCAGGTCTCTCTCTGACTACTCTCCAGAAATTTTC	2073	634	GAAGAAACCCATTTGGAAACAGTCCGAAATTCCTTCCATGATATACCAAGGAAGTGA 693
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620	141	GluArgLysArgIleGlyThrValGlyLysLeuAspTyrAspTyrHisGlnGlySerGly 160
Qy	2074	TCITTTCAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTA	2133	694	ACATTTCTGTTTCAAGCCGTAGTGGAAATTTATCACGTAAAGATGAGAGGCCACAGGA 753
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyLysProGlnGly 180
Qy	2134	CAGCTCGAAGAAATCTCTACTGTGCTGTTCATATATGFGGT 2178		754	TTTACGCAACACCTTTTMAGCCCAATCTAGTGAACACTAGTTGTCCCAACATACCGATG 813
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655		181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
RESULT 6					
; Sequence 21, Application US/09976674					
; Patent No. 6844180					
; GENERAL INFORMATION:					
; APPLICANT: Qi, Steve					
; APPLICANT: Akinsanya, Karen					
; APPLICANT: Riviere, Pierre					
; APPLICANT: Junien, Jean-Louis					
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV					
; FILE REFERENCE: 70669					
; CURRENT APPLICATION NUMBER: US/09/976,674					
; CURRENT FILING DATE: 2001-10-12					
; PRIOR FILING DATE: 2001-10-12					
; PRIOR FILING DATE: 2000-10-12					
; NUMBER OF SEQ ID NOS: 61					
; SOFTWARE: Patent in version 3.1					
; SEQ ID NO 21					
; LENGTH: 613					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-09-976-674-21					
Alignment Scores:					
Pred. No.:	0	Length:	613		
Score:	3236.00	Matches:	607		
Percent Similarity:	100.0%	Conservative:	0		
Best Local Similarity:	100.0%	Mismatches:	0		
Query Match:	58.3%	Indels:	0		
DB:	2	Gaps:	0		

```

Db      341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy      1294 TATATTGCCAGAGCTGGATGACTCTCTGAGGGAAATATGCTTGGTCCATCTACTAGAT 1353
Db      361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Qy      1354 CGTCCCAGAGCTCGCTCAGACAGTAGTGTTGATCTCACCTGAATATTATTATCCAGTAGAA 1413
Db      381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Qy      1414 GATGATCTTATGAAAGGACAGACTCATTGAGTCAGTGCTCTGATTTCTGTGACGCCACTA 1473
Db      401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Qy      1474 ATTATCTATCAAGAAACAACAGACATCTCGATAAATATCCATGACATCTTTCATGTTTTT 1533
Db      421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
Qy      1534 CCCAAAGTCACGAAGAGAAATGAGTTTATTTTTGCTCTGGAATGCAAAACAGGTTTC 1593
Db      441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Qy      1594 CGTCATTTATCAAAATACATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGT 1653
Db      461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Qy      1654 GGGCTGCTCTCCNAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTACCAGT 1713
Db      481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Qy      1714 GETCAATGGGAAGTCTTGGCCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAAG 1773
Db      501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Qy      1774 CTGTATATTTTGAAGCACCAAGACTCCCTCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db      521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
Qy      1834 TACGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACATCTTGTGTCG 1893
Db      541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Qy      1894 ATCAGTCAGACTGTGACTTCTTTATTAAGTATAGTAAACAGAGAGATCCACACTGT 1953
Db      561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy      1954 GTGTCCCTTACAGACTATCAAGTCTCTGAGATGCCCAACTTGCACAAACAAAGGAATTT 2013
Db      581 ValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPhe 600
Qy      2014 TGGGCCACCATTTTGGATTCA 2034
Db      601 TrpAlaThrIleLeuAspSer 607

RESULT 7
US-09-976-674-23
; Sequence 23, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Oi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976.674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 892
;
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-23
Alignment Scores:
Pred. No.: 9,06e-296 Length: 892
Score: 2871.00 Matches: 529
Percent Similarity: 73.7% Conservative: 137
Best Local Similarity: 58.5% Mismatches: 208
Query Match: 51.7% Indels: 30
DB: 2 Gaps: 5
US-10-825-632-2 (1-3120) x US-09-976-674-23 (1-892)
Qy      124 TGGAGGCGGCGCAGCATGAAGCGGCGCAGCGCTCCATAGCGCAGTCGCGAGCGGTCC 183
Db      17 TrpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
Qy      184 GGGCGGGCGCGGGGGAAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACAGCTG 243
Db      37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45
Qy      244 CGTGTGTGAGATATTTGAAACTGCGGACTGCGAGGAGATATTGAATCACAGGATCGCGCT 303
Db      46 -----ThrAspAspPro 49
Qy      304 AAATTGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTTGCC 363
Db      50 AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis 68
Qy      364 GATACAGAAATATCATCGCTACATGATGCTAAGGCACACACATGATTTTCATGTTTGTG 423
Db      69 GlySerArgLysTyrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal 88
Qy      424 AAGAGGATGATCCAGATGGACCTCATTCACAGAGATCTATTACCTTGCATCTCTGGT 483
Db      89 GlnLysThrAspGluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyr 108
Qy      484 GAGAACAGAGAAATATACACTGTTTATTCTGAAATTCACAAACTATCAATAGACAGCA 543
Db      109 GlySerArgGluAsnSerLeuLeuTyrSerGluIleProLysLysValArgLysGluAla 128
Qy      544 GTCTTAATCTCTCTTGAAGCTCTTTTGGATCTTTTTCAGCAACATCGACTATGGA 603
Db      129 LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisGly 148
Qy      604 ATGTATTCGAGAGAGAACTATTAAAGAGAAAGAAACGCATTGGACACAGTCGGATT 663
Db      149 ValTyrSerArgGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle 168
Qy      664 GCTTCTTACGATTATCACCAGAAAGTGGACATTTCTTCTTCAAGCCGGTAGTGAATT 723
Db      169 ThrSerTyrAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeu 188
Qy      724 TATCAGTAAAGATGGAGGCCACAGGATTTACGCAACACTTTAAGCCCAATCTA 783
Db      189 PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGlu 208
Qy      784 GTGGAACACTAGTTGTCCCAACATACGATGGATCCAAATATTGCCCGCTGATCCAGAC 843
Db      209 IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla 228
Qy      844 TGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTGTAACACAGAGAA 903
Db      229 PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGlu 248
Qy      904 AGGAGACTCATCTTATGTGCACATAGCTAGCCCAACATGGAAAGAGATGCCAGATCAGCT 963
Db      249 ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAla 268
Qy      964 GGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGTATTCTGCTATTGCTGGTGT 1023
Db      269 GlyValAlaThrPheValIleGlnGluGluPheAspArgPheThrGlyTyrTrpTrpCys 288
```



```
QY 1024 CCAAAAGCTGAACAACTCCAGTGGTGT---AAATTTCTAGAAATCTATATGAGAA 1080
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 289 ProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgGileuLysTrpGluGlu 308
QY 1081 AATGATGAATCTCAGGTGGAAATATTATCATGTTTACATCCCTCATGTTTGGAAACAGAGG 1140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 309 ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluGluGluArgLys 328
QY 1141 GCAGATTTCATTCGTTTATCTCTAAACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGAT 1200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 329 ThrAspSerTyrArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu 348
QY 1201 TCAGAAATAATATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACATAATT 1260
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 349 AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal 368
QY 1261 CAACCTTTTCAGATTCTATTGGAAGGATTGAAATATATTGCCAGAGCTGGATGCACTCCT 1320
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 369 GlnProPheSerSerLeuPheProLysValGluTyrIleAlaArgAlaGlyTrpThrArg 388
QY 1321 GAGGGAATAATATCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTCACATAGTG 1380
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 389 AspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal 408
QY 1381 TTGATCTCACTGAAATATTATCCAGTAGAAGATGATGTTATCGAAAGCAGAGACTC 1440
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 409 LeuLeuProProAlaLeuPheIleProSerThrGluAsnGluGluGlnArgLeuAlaSer 428
QY 1441 ATTGAGTCAGTGCCTGATCTGTGACGCCACTAATATTCTATGAAGAAACACAGACATC 1500
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 429 AlaArgAlaValProArgAsnValGlnProTyrValValTyrValTyrValThrAsnVal 448
QY 1501 TGGATAAATATCCATGATCTTTTCATGTTTTCCTCCAAAGTCAC---GAAGAGAAAT 1557
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 449 TrpIleAsnValHisAspIlePheTyrProPheProGlnSerGluGlyGluAspGluLeu 468
QY 1558 GAGTTTATTTTTCCTCTGAATCAAAACAGGTTTCCGTCATTTATACAAATATACATCT 1617
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 469 CysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHisLeuTyrLysValThrAla 488
QY 1618 ATTTTAAAGGAAAGCAATATAAACAATCCAGTGGTGGCTCCCTGCTCCCAAGTGATTC 1677
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 489 ValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSerProGlyGluAspGluPhe 508
QY 1678 AAGTGTCTCTATCAAGAGGAGATAGCAATTACCAGTGGTGAATGGAAAGTTCTTGGCGG 1737
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 509 LysCysProIleLysGluGluIleAlaLeuThrSerGlyLysValLeuAlaArg 528
QY 1738 CATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGTATATTTTGAAGGCACCAAA 1797
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 529 HisGlySerLysIleTrpValAsnGluThrLysLeuValTyrPheGlnGlyThrLys 548
QY 1798 GACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAGTAATCCCTGGAGAGTGACA 1857
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 549 AspThrProLeuGluHisLeuTyrValValSerTyrGluAlaAlaGlyGluIleVal 568
QY 1858 AGGTGACTGACCGTGGCTACTACATCTTCTGTCGATCAGTCAGTCAGTGTCTCTTT 1917
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 569 ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPhe 588
QY 1918 ATAAGTAAGTATAGTAACAGAGAAGATCCACACTGTGTCTCCCTTTACAGCTATCAAGT 1977
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 589 ValSerHisTyrSerSerValSerThrProProCysValHisValTyrLysLeuSerGly 608
QY 1978 CCTGAAGATGACCCCACTTCGAAACAAAGGAATTTTGGCCACCACCTTTTGATTCAGCA 2037
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 609 ProAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAla 628
QY 2038 GGTCTCTCTGACTATACCTCCAGAAATTTTCTCTTTTGAAGAACTACTGATTT 2097
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 629 SerCysProProAspTyrValProProGluIlePheHisPheHisThrArgSerAspVal 648
```

RESULT 8

```
US-09-976-674-27
; Sequence 27, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Jumeire, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 27
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
```


US-09-976-674-27

Alignment Scores:

Pred. No.: 9,06e-296 Length: 892
Score: 2871.00 Matches: 529
Percent Similarity: 73.7% Conservative: 137
Best Local Similarity: 58.5% Mismatches: 208
Query Match: 51.7% Indels: 30
DB: 2 Gaps: 5

US-10-825-632-2 (1-3120) x US-09-976-674-27 (1-892)

QY	124	TGGAGCGCGGCAGCATGAACGGCGCGAGCGCCGCTCCATAGCGCAGCTCGGACGGTCC	183
DB	17	TpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro	36
QY	184	GGGCGGGCGGGGGAAGGAATGCACATGCGCAGCAGCAATGGAACAGACAGCTG	243
DB	37	ThrAlaAspA--GGlyAsp-----AlaAlaAla-----	45
QY	244	GGTGTGTAGATATTGAAACTGCGGACTGTGAGGAGAATATTGAATCACAGGATCGGCT	303
DB	46	-----ThrAspAspPro	49
QY	304	AAATTGGAGCCTTTTATGTGAGCGGTATTCTCTGGAGTCAGCTTAAAGCTGCTGCC	363
DB	50	AlaAlaArg--PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis	68
QY	364	GATACCAAGAAATATCATGTGTACATGATGCTTAAGGCACACATGATTTTCATGTTGG	423
DB	69	GlySerArgLysTySerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal	88
QY	424	AAGAGGATGATCCAGATGAGACCTTATTCAGACAGATCTATTACCTTGGCCTGCTGT	483
DB	89	GlnLysThrAspGluSerGlyProHisSerHisArgLeuTyTrpLeuGlyMetProTyTr	108
QY	484	GAGACAGAGAAATACACTGTTTATCTGAATTCCTCAAAATCCAAACTATCAATAGCAGCA	543
DB	109	GlySerArgGluAsnSerLeuLeuTySerGluIleProLysLysValArgLysGluAla	128
QY	544	GTCTTAAGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGA	603
DB	129	LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisHisGly	148
QY	604	ATGTATCTCGAGAGAGAACTATTAAAGAGAAAGAAACCATTTGGAACAGTCGGAATT	663
DB	149	ValTySerArgGluGluGluLeuLeuArgLysArgLeuGlyValPheGlyIle	168
QY	664	GCTTCTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGAATT	723
DB	169	ThrSerTyAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeu	188
QY	724	TATCAGCTAAAGATGAGGCGCCACAGGATTTACGCAACAACCTTTAAGGCCCAATCTA	783
DB	189	PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGlu	208
QY	784	GTGGAACATGATGTTGCCCAACATCAGGATGATCAAAATATGCCCGCTGATCCAGAC	843
DB	209	IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla	228
QY	844	TGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	903
DB	229	PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGlu	248
QY	904	AGGAGACTCATTATGTGCACAAATGAGCTAGCCAACTAGGAAGAGATGCCAGATCAGCT	963
DB	249	ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAla	268
QY	964	GGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGTGT	1023
DB	269	GlyValAlaThrPheValIleGlnGluPheAspArgPheThrGlyTyTrpTrpCys	288
QY	1024	CCAAAAGCTGAAACAACTCCACAGTGGTGGT---AAAATCTTAGAATTCATATGAGAA	1080

DB	289	ProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyTrpGluGlu	308
QY	1081	AATGATGAATCTGAGTGGAAATATTATTCATGCTTACATCCCTATGTTGGAAACAGAGGAG	1140
DB	309	ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluGluArgLys	328
QY	1141	GCAGATTTCCTGCTTATCCTAAACACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATG	1200
DB	329	ThrAspSerTyArgTyProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu	348
QY	1201	TCAGAAATATGATGCTGAGGAAAGGATCATAGATCATAGATAGTAAAGAACTAAT	1260
DB	349	AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal	368
QY	1261	CAACCTTTTGAGATCTTATTGAAGAGTTGAATATATTGCCAGAGCTGGATGACTCCT	1320
DB	369	GlnProPheSerSerLeuPheProLysValGluTyIleAlaArgAlaGlyTrpThrArg	388
QY	1321	GAGGAAATATGCTTGTGCTCCTACTAGATCGCTCCAGACTCGCTACATAGTAGT	1380
DB	389	AspGlyLysTyAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal	408
QY	1381	TTGATCTCACCTGAATTTATTCCTCAGTAGAGATGATGTTATGGAAGGCACAGACTC	1440
DB	409	LeuLeuProProAlaLeuPheIleProSerThrGlnAsnGluGlnArgLeuAlaSer	428
QY	1441	ATTGAGTCACTGCTGATTCCTGAGCGGCACCTAATTTATCTATGAAGAAACACAGACATC	1500
DB	429	AlaArgAlaValProArgAsnValGlnProTyValValTyGluGluValThrAsnVal	448
QY	1501	TGATATAATATCCATGACATCTTTCATGTTTTCCTCCCAAGTCAC---GAAGAGGAAAT	1557
DB	449	TrpIleAsnValHisAspIlePheTyProPheProGlnSerGluGlyGluAspGluLeu	468
QY	1558	GAGTTATTTTTCCTCTGAATGCAAAACAGAGTTTTCCTCATTTATACAAAATTCATCT	1617
DB	469	CysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHisLeuTyLysValThrAla	488
QY	1618	ATTTTAAAGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTCCCAAGTGATTC	1677
DB	489	ValLeuLysSerGlnGlyTyAspTrpSerGluProPheSerProGlyGluAspGluPhe	508
QY	1678	AAGTGTCTTATCAAGAGGAGATAGCAATTACCAGTGGTGAATGGAGATCTTCTGGCGG	1737
DB	509	LysCysProIleLysGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArg	528
QY	1738	CATGGATCTAATATCAAGTTGATGATGAGTCAGAGGCTGTTATTTTGAAGGCACCAAA	1797
DB	529	HisGlySerLysIleTrpValAsnGluGluThrLysLeuValTyPheGlnGlyThrLys	548
QY	1798	GACTCCCTTTAGAGCATCACCTGTAGTAGTACGTACGTAAATCTCGGAGAGGTGACA	1857
DB	549	AspThrProLeuGluHisIleLeuTyValValSerTyGluAlaAlaGlyGluIleVal	568
QY	1858	AGGTGACTGACCGCTGCTACTCACATCTTCTGCTGCATCAGTCAGCTGCTGCTCTTT	1917
DB	569	ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPhe	588
QY	1918	ATAAGTAAGTATAGTAAACCAGAGAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGT	1977
DB	589	ValSerHisTySerSerValSerThrProCysValHisValTyLysLeuSerGly	608
QY	1978	CCTGAAGATGACCAACTTGCACAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCA	2037
DB	609	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAla	628
QY	2038	GGTCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTT	2097
DB	629	SerCysProProAspTyValProProGluIlePheHisPheHisThrArgSerAspVal	648
QY	2098	ACATTGTATGGGATGCTTACAGGCTCATGATCTACAGCTCGGAGCAATATCTACT	2157

Db	649	ArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThr	668
Qy	2158	GTGCTGTCATATATGGTGGTCTCCAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTC	2217
Db	669	ValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyLe	688
Qy	2218	AAGTATTTCGGCTTGAATACCCCTAGCCTCTCTAGGTATGTGGTTGTAGTATAGACAAC	2277
Db	689	LysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGly	708
Qy	2278	AGGGATCCCTGCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAAATGGGTCAA	2337
Db	709	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGln	728
Qy	2338	ATAGAAATTCAGCATCAGGTGGAAGCACTCCATATCTAGCTTCTCGATATATGATTTCAAT	2397
Db	729	ValGluLeuGlnAspGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIle	748
Qy	2398	GACTAGATCGTGTGGGCATCCAGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGCCA	2457
Db	749	AspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGlyPheLeuSerLeuMetGly	768
Qy	2458	TTAATGCAGAGTCAGATATCTTCAGGTTTCTATGCTGGGGCCCGCCAGTCACTCTGTGG	2517
Db	769	LeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAlaProValThrValTyr	788
Qy	2518	ATCTTCATGATACAGGATACAGGAACGTTATATGGTTCACCCCTGACCAAGATGACAG	2577
Db	789	MetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHis	808
Qy	2578	GGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAATTCCTCTGACCAAAATCGT	2637
Db	809	GlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuProAsnGlnProAsnArg	828
Qy	2638	TTACTGCTCTTACATGTTTCTCTGGATGAGATGTCCATTTTGCACATACCATATATTA	2697
Db	829	LeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeu	848
Qy	2698	CTGAGTCTTTTGTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGA	2757
Db	849	ValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArg	868
Qy	2758	CACAGCATAAGAGTTCTCTGAATCGGAGAACATTTATGAAGTGCATCTTTTGCACATACCTT	2817
Db	869	HisSerIleArgCysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeu	888
Qy	2818	CAGGAACCTT 2829	
Db	889	GlnGluTyrLeu 892	
RESULT 9			
US-09-976-674-3			
; Sequence 3, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 3			
; LENGTH: 863			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-976-674-3			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-10-825-632-2 (1-3120) x US-09-976-674-3 (1-863)			
Qy	316	TTTATGTTGACCGGTATCTCTGAGTCACGCTTAAAGCTGCTTCCCATACCAGAAA	375
Db	24	PheGlnValGlnLysHisSerTyrAspGlyLeuArgSerIleIleHisGlySerArgLys	43
Qy	376	TATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTTGTGAAGGAATGAT	435
Db	44	TyrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheValGlnLysThrAsp	63
Qy	436	CCAGATGCACCTCAATTCAGACAGAAATCTATTACCTTCCCATGCTCTCGTGAGAACAGAA	495
Db	64	GluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyrGlySerArgGlu	83
Qy	496	AATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGACGACAGCTTCTTAATGCTC	555
Db	84	AsnSerLeuLeuTyrSerGluIleProLysLysValArgLysGluAlaLeuLeuLeu	103
Qy	556	TCCTCGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTATGGAATGTATTCTCGA	615
Db	104	SerTyrLysGlnMetLeuAspHisPheGlnAlaThrProHisGlyValTyrSerArg	123
Qy	616	GAAGAAGAACTATTAAAGAGAAGAAACGATTTGGAACAGCTCGGAATTTGCTTACGAT	675
Db	124	GluGluGluLeuLeuLeuArgLysArgLeuGlyValPheGlyIleThrSerTyrAsp	143
Qy	676	TATCACCAAGNAGTGGNACATTTCTGTTTCACGCGGTAGTGGATTTATACGTAATAA	735
Db	144	PheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheHisCysArg	163
Qy	736	GATGAGGCCCAAGAGTTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGT	795
Db	164	AspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGluIleLysThrGln	183
Qy	796	TGTCCCAACATACGATCGATCCAAAATTTATGCCCTGATCCCTGATCCAGACTGGATTCT	855
Db	184	CysSerGlyProArgMetAspProLysIleCysProAlaAspProAlaPheSerPhe	203
Qy	856	ATACATAGCACGATATTGGATATCTTAACATCTGTAAACAGAGAGAAAGAGACTCACT	915
Db	204	IleAsnAsnSerAspLeuTyrPheValAlaAsnIleGluThrGlyGluGluArgGluThr	223
Qy	916	TATGTGCAATGAGCTAGCCCAACATGGAAGAAGATCCAGATCAGCTGAGTGCCTTACC	975
Db	224	PheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAlaGlyValAlaThr	243
Qy	976	TTTGTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTCTGGCTGCTGCCAAAGCTGAA	1035
Db	244	PheValIleGlnGluGluPheAspArgPheThrGlyTyrTrpTyrCysProThrAlaSer	263
Qy	1036	ACAACTCCAGTGGTGGT---AAAATTTCTTAAATTTCTATATGAAGAAAATGATGATCT	1092
Db	264	TrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGluValAspGluSer	283
Qy	1093	GAGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACAAAGAGGCGGAGATTCTATC	1152
Db	284	GluValGluValIleHisValProSerProAlaLeuGluArgLysThrAspSerTyr	303
Qy	1153	CGTATCTCTAAACAGGTACACAAATCTCAAGTCACTTTTAAAGTGTCTGAAATAATG	1212
Db	304	ArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGln	323
Qy	1213	ATTGATGCTGAAGGAAGATCATAGATGCTCATAGTAAGCACTAATTTCACTTTGAG	1272
Db	324	ThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSer	343

```
QY 1273 ATTCTATTGAAGGAGTTGAATATATTGCGAGAGCTGGATGGACTCCTCGAGGAAATAT 1332
Db |||||||
QY 344 SerLeuPheProLysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyr 363
Db |||||||
QY 1333 GCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTATCTCACCT 1392
Db |||||||
QY 364 AlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuValLeuLeuProPro 383
Db |||||||
QY 1393 GAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGAGGAGAGACTCATTTGAGTCAGTG 1452
Db |||||||
QY 384 AlaLeuPheIleProSerThrGluAsnGluGluArgLeuAlaSerAlaArgAlaVal 403
Db |||||||
QY 1453 CCTGATCTCTGAGCCCACTAATTTATCTATGAGAACAACACACATCTGGTAATATC 1512
Db |||||||
QY 404 ProArgAsnValGlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnVal 423
Db |||||||
QY 1513 CATGACATCTTTTCATGTTTTCCTCCCAAGTCAC---GAAGAGGAATTCAGTTTATTTT 1569
Db |||||||
QY 424 HisAspIlePheTyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArg 443
Db |||||||
QY 1570 GCCTCTGAATGCAAAACAGGTTTCCGTCAATTATACAAATTAACATCTATTTTAAAGGAA 1629
Db |||||||
QY 444 AlaAsnGluCysLysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSer 463
Db |||||||
QY 1630 AGCAATATAAACCATCCAGTCGTGGCTGCTCCAGTCAATTCAGTGTCTATC 1689
Db |||||||
QY 464 GlnGlyTyrAspTrpSerGluProPheSerProGlyGluAspLupheLysCysProIle 483
Db |||||||
QY 1690 AAAGAGGAGATAGCAATTTACAGTGTGTAATGGGAAGTCTTGGCCGGCATGATCTAAT 1749
Db |||||||
QY 484 LysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLys 503
Db |||||||
QY 1750 ATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTA 1809
Db |||||||
QY 504 IleTrpValAsnGluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeu 523
Db |||||||
QY 1810 GAGCATACCTGTAGTAGTCAGTTAGTAAATCCCTGGAGAGGTGACAGGCTGACTGAC 1869
Db |||||||
QY 524 GluHisHisLeuTyrValValSerTyrGluAlaAlaGlyLupheValArgLeuThrThr 543
Db |||||||
QY 1870 CGTGGCTACTCACATCTCTGTCATCAGTCAGTCAGTCTCTTTATTAAGTAAGTAT 1929
Db |||||||
QY 544 ProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyr 563
Db |||||||
QY 1930 AGTAACCAAGAAGATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCTGAAGATGAC 1989
Db |||||||
QY 564 SerSerValSerThrProProCysValHisValTyrLysLeuSerGlyProAspAspAsp 583
Db |||||||
QY 1990 CCAACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCT 2049
Db |||||||
QY 584 ProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProPro 603
Db |||||||
QY 2050 GACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTACATGTTATGGG 2109
Db |||||||
QY 604 AspTyrValProProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGly 623
Db |||||||
QY 2110 ATGCTCTACAGCCTCATGATCTACAGCTGGGAAGAAATATCTTACTGTGTTTCATA 2169
Db |||||||
QY 624 MetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheVal 643
Db |||||||
QY 2170 TATGCTGTCTCTCAGTGCAGTGTGTAATTCGGTTTAAAGGAGTCAAGTATTTCCGC 2229
Db |||||||
QY 644 TyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArg 663
Db |||||||
QY 2230 TTGAATACCTAGCTCTCTAGGTTATGCTGTTGCTAGTAGACAAACAGGAGGATCCTGT 2289
Db |||||||
QY 664 LeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCys 683
Db |||||||
QY 2290 CACCGAGGCGCTTAAATTTGAAGGCGCCTTTTAAATATATAAATGGGTCAAAATGAC 2349
Db |||||||
QY 684 GlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGlu 703
Db |||||||
```

```
QY 2350 GATCAGGTGGAAGGACTCCAATATATCTAGCTTTCTCGATATATGTTTCATTGATTCGT 2409
Db |||||||
QY 704 AspGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArg 723
Db |||||||
QY 2410 GTGGGCAATCCAGCGTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATCCAGAGG 2469
Db |||||||
QY 724 ValAlaIleHisGlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLys 743
Db |||||||
QY 2470 TCAGATATCTTCAGGGTTCCTATTCCTGGGGCCCGAGTCACCTCTGTGATCTTCTATGAT 2529
Db |||||||
QY 744 ProGlnValPheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaIatyrAsp 763
Db |||||||
QY 2530 ACAGGATACACGAAGCTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTA 2589
Db |||||||
QY 764 ThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAla 783
Db |||||||
QY 2590 GGATCTGTGGCCATCAGACAGAAAGTTCCCTCTGTAACCAATCGTTTACTGCTCTTA 2649
Db |||||||
QY 784 GlySerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleLeu 803
Db |||||||
QY 2650 CATGGTTTCTCGATGAGATGTCCATTTTGGACATACCATGATATTTACTGAGTTTTTTA 2709
Db |||||||
QY 804 HisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeu 823
Db |||||||
QY 2710 GTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACACAGCATAGA 2769
Db |||||||
QY 824 IleArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArg 843
Db |||||||
QY 2770 GTTCTCTGAATCGGAGAACATTTATGACTGTCATCTTTTGGCACTACCTTCAAGAAACCTT 2829
Db |||||||
QY 844 CysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 863
Db |||||||

RESULT 10
US-09-976-674-33
; Sequence 33, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCES: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-33

Alignment Scores:
Pred. No.: 1,66e-290 Length: 879
Score: 2821.50 Matches: 522
Percent Similarity: 72.7% Conservative: 135
Best Local Similarity: 57.7% Mismatches: 204
Query Match: 50.8% Indels: 43
DB: 2 Gaps: 6

US-10-825-632-2 (1-3120) x US-09-976-674-33 (1-879)
QY 124 TGGAGGCGGCGCAGCATGTAAGCGCGCAGCGCCGCTCCATAGCCGACGTCGGGACGTC 183
Db |||||||
QY 17 TrpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
Db |||||||
QY 184 GGC GCGGCGCGCGGGAAGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTG 243
Db |||||||
QY 37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45
```

QY	244	GGTGTGAGATATTTGAAACTCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCT	303	Db	389	AspGlySerValAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal	408
Db	46	-----ThrAspAspPro	49	QY	1381	TTGATCTCACTGATATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGAGACTC	1440
QY	304	AAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTCTTGC	363	Db	409	LeuLeuProProAlaLeuPheIleProSerThrGluAsnGluGluGlnArgLeuAlaSer	428
Db	50	AlaAlaArg---PheGlnValGlnIleHisSerTrpAspGlyLeuArgSerIleHis	68	QY	1441	ATTGAGTCAGTCCTGATCTCTGTGACGCCACTAATTTATCTATGAGAAGAACACACAGCATC	1500
QY	364	GATACACAGAAATATCATCGCTACATGATGGCTAAGGACACACATGATTTTCATGTTGTG	423	Db	429	AlaArgAlaValProArgAsnValGlnProTyrValValTyrGluGluValThrAsnVal	448
Db	69	GlySerArgIleTyrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal	88	QY	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCAC---GAAGAGGAAAT	1557
QY	424	AAGAGGAATATCCAGATCGACTCAATTCAGACAGAACTATTACCTTGGCCATCTCGGT	483	Db	449	TrpIleAsnValHisAspIlePheTyrProPheProGlnSerGluGlyGluAspGluLeu	468
Db	89	GlnLysThrAspGluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyr	108	QY	1558	GAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTCATCT	1617
QY	484	GAGAACAGAGAAATATACATGTTTATCTGAAATTCGAAATCCAAACTATCAATAGACAGCA	543	Db	469	CysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHisLeuTyrLysValThrAla	488
Db	109	GlySerArgGluAsnSerLeuLeuTyrSerGluIleProTyrLysValArgLysGluAla	128	QY	1618	ATTTTAAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTCCAGTGATTTTC	1677
QY	544	GTCTTAATGCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACTCGACTGTGGA	603	Db	489	ValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSerProGlyGluAspGluPhe	508
Db	129	LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisGly	148	QY	1678	AGTGTCTCTATCAAGAGGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGG	1737
QY	604	ATGTATTCTCGAGAAGAAGAACTATTAAAGAGAAAGAAACGCATTTGGAAACAGTCGGAAT	663	Db	509	LysCysProIleLysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArg	528
Db	149	ValTyrSerArgGluGluGluLeuLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle	168	QY	1738	CATGATCTAATATCAAGTTGATGAAGTCAGAGGCTGCTATATTTTGAAGGCCACCAAA	1797
QY	664	GCTTCTTACGATATACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGCTAGTGGAAAT	723	Db	529	HisGlySer-----LysGlyThrLys	535
Db	169	ThrSerTyrAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeu	188	QY	1798	GACTCCCTTTAGAGCATCCTGTAGTAGTCAGTTACGTAAATCCTCGAGAGGTGACA	1857
QY	724	TATCACGTAAAGATGAGGCGCCACAGGATTTTACGCAACACCTTTAAGGCCCAATCTA	783	Db	536	AspThrProLeuGluHisLeuTyrValSerTyrGluAlaGlyGluIleVal	555
Db	189	PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGlu	208	QY	1858	AGGCTGACTGACCGTGGCTACTACATCTTGTGTCAGTCAGTCAGCAGCTGTGACTCTTT	1917
QY	784	GTGGAATCTGTTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCAGAC	843	Db	556	ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPhe	575
Db	209	IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla	228	QY	1918	ATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGT	1977
QY	844	TGGATTGCTTTTATACATACACAGATTTTGGATATCTAAATCTGTAACCAAGAGAA	903	Db	576	ValSerHisTyrSerSerValSerThrProCysValHisValTyrLysLeuSerGly	595
Db	229	PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGlu	248	QY	1978	CCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTTTGGGCCACCATTTTGGATTCAGCA	2037
QY	904	AGGAGACTCATTTATGTGCACATGACCTAGCCCAACATGGAAGATGCCAGATCAGCT	963	Db	596	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetGluAlaAla	615
Db	249	ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspPheProLysSerAla	268	QY	2038	GGTCTCTTCTCGACTATACCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTT	2097
QY	964	GGAGTCGTACTTTTCTCCAAAGAAATTTGATAGATATTCGCTATTTCGTGTGT	1023	Db	616	SerCysProProAspTyrValProProGluIlePheHisPheHisThrArgSerAspVal	635
Db	269	GlyValAlaThrPheValIleGlnGluGluPheAspArgPheThrGlyTyrTrpCys	288	QY	2098	ACATTTGATGGGATGCTCTACAAAGCTCATGATCTACAGCTCGAAAGAAATATCCTACT	2157
QY	1024	CCAAAAGCTGAACAACTCCAGTGGTGGT---AAAATCTTAGAATCTCTATGAGAA	1080	Db	636	ArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThr	655
Db	289	ProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGlu	308	QY	2158	GTGCTGTTTCATATATGTTGGTCTCAGCTGAGTTGGTGAATAATCGTTTAAAGGAGTC	2217
QY	1081	AATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAACAGAGG	1140	Db	656	ValLeuPheValTyrGlyProGlnValGlnLeuValAsnSerPheLysGlyIle	675
Db	309	ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluArgLys	328	QY	2218	AAGTATTTCCGTTTCATACCTAGCTCTCTAGTATTGTTGTTGATGATGACAGAAC	2277
QY	1141	GCAGATTCAATTCGTTATCTTAAACAGGTACAGCAAAATCCTAAAGTCATTTTAAAGATG	1200	Db	676	LysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGly	695
Db	329	ThrAspSerTyrArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu	348	QY	2278	AGGGGATCTCTGTCCAGCGGCTTAAATTTGAGCGGCTTTAAATATATAATATGATTCAT	2337
QY	1201	TCAGAAATAATGATTGATGCTGAAGAGGATCATAGATGCTATAGATAAGGAACATAAT	1260	Db	696	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGln	715
Db	349	AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal	368	QY	2338	ATAGAAATTCAGCATCAGGTGGAAGACTCCCATATCTAGCTTCTCGATGATTCAT	2397
QY	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTCGAGAGCTGATGAGCTCCT	1320	Db	716	ValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIle	735
Db	369	GlnProPheSerSerLeuPhePheProLysValGluTyrIleAlaArgAlaGlyTrpThrArg	388	QY	2398	GACTTAGATCGTGTGGGCATCCAGCGCTGCTCTATGAGGAGTACCTCTCCCTGATGCA	2457
QY	1321	GAGGGAATAATGCTTGGTCCATCCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTG	1380				

Db 736 AspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPheLeuSerLeuMetGly 755
Qy 2458 TTAATGAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCCAGTCACTCTGGG 2517
Db 756 LeuIleHisLysProGlnValPheLeValAlaIleAlaGlyAlaProValThrValTrp 775
Qy 2518 ATCTTCATATACAGGATACAGGAACGTTATATGGGTCAACCCTGACACAGAAACAG 2577
Db 776 MetAlaIyrAspThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHis 795
Qy 2578 GGCTATTACTTAGGATCTGGCCATGCAAGCAGAAAAGTCCCTCTGAACCAATCGT 2637
Db 796 GlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuProAsnGluProAsnArg 815
Qy 2638 TTAAGTCTTACATGATGTTCTGGATGAGAATGTCATTTGGCATACATACAGATATATA 2697
Db 816 LeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeu 835
Qy 2698 CTGAGTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGA 2757
Db 836 ValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArg 855
Qy 2758 CACAGCATAGAGTTCTGAAATCGGAGAACATTAATGAATGCTATCTTTTGCACTACCTT 2817
Db 856 HisSerIleArgCysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeu 875
Qy 2818 CAAGAAAACCTT 2829
Db 876 GlnGluTyrLeu 879

RESULT 11
US-09-976-674-35
; Sequence 35, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-35

Alignment Scores:
Pred. No.: 1.66e-290 Length: 879
Score: 2821.50 Matches: 522
Percent Similarity: 72.7% Conservative: 135
Best Local Similarity: 57.8% Mismatches: 204
Query Match: 50.8% Indels: 43
DB: 2 Gaps: 6

US-10-825-632-2 (1-3120) x US-09-976-674-35 (1-879)

Qy 124 TGGAGCGCGCAGCATGTAACGGCGGAGCCCGCTCCATAGCGCAGCTCGGACGGTCC 183
Db 17 TrpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
Qy 184 GGGCGGGCGGGGGGGAAGAAATGCAATCGCAGCAGCAATGGAACAGACAGCTG 243
Db 37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45
Qy 244 GGTGTTGAGATATTGAAACTCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCT 303

46 -----ThrAspAspPro 49
304 AAATTGGACGCTTTTATGTTGAGCGTATTCCTGGAGTCAGCTTAAAGCTGCTTGCC 363
Db 50 AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis 68
Qy 364 GATACACAGAAAATATCATGCTTACATGATGAAGCCACCATGATGATTGATGTTGTG 423
Db 69 GlySerArgLysTyrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal 88
Qy 424 AAGAGGAATGATCCAGATGGAGCTCATTCAGACAGAAATCTATTACCTTGCCATGCTCGT 483
Db 89 GlnLysThrAspGluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyr 108
Qy 484 GAGAACAGAGAAAATACACTGTTTATTCGAAATTCACAAACTATATCAATAGACAGCA 543
Db 109 GlySerArgGluAsnSerLeuLeuTyrSerGluIleProLysLysValArgLysGluAla 128
Qy 544 GTCTTAATCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGA 603
Db 129 LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisGly 148
Qy 604 ATGTATTCTCGAGAGAAAGAACTATTAAAGAGAGAAAGAAACGATTTGGAACAGTCGGAATT 663
Db 149 ValTyrSerArgGluGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle 168
Qy 664 GCTTCTTACGATTTATCACCAAGAAAGTGAACATTTCTGTTTCAAGCGGTGATGGAATT 723
Db 169 ThrSerTyrAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeu 188
Qy 724 TATCAGTAAAGATGGAGGCCACAGGATTTACGCAACCACTTTAAGGCCCAATCTA 783
Db 189 PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysLeuGlu 208
Qy 784 GTGGAACACTAGTTGTCCCAACATACGATGGATGATCAAAATTTATGCCCGGTGATCCAGAC 843
Db 209 IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla 228
Qy 844 TGGATTGCTTTTATACATAGCAACGATTTTGGATATCTAAACATCGTAAACAGAGAA 903
Db 229 PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGlu 248
Qy 904 AGGAGACTCACTTGTGCAATGAGCTAGCCATGCAACATGGAAGAGATGCCAGATCAGCT 963
Db 249 ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAla 268
Qy 964 GGAGTCGCTACCTTTGTTCTCCAAGAAGAATTTGATAGATATCTGGCTATTGTTGTGT 1023
Db 269 GlyValAlaThrPheValIleGlnGluPheAspArgPheThrGlyTyrTrpTrpCys 288
Qy 1024 CCAAAAGCTGAACAACTCCAGTGGTGT---AAAATCTAGNATTCATATATGAAGAA 1080
Db 289 ProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGlu 308
Qy 1081 AATGATGAATCTGAGGTGGAATTTATCATGTTATCATCCCTATGTTGGAACAGAGG 1140
Db 309 ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluGluArgLys 328
Qy 1141 GCAGATTTCATTCGTTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGAT 1200
Db 329 ThrAspSerTyrArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu 348
Qy 1201 TCAGAAATATGATGATGCTGAAGAGGATCATAGATGTGATAGATAAGAACTAATT 1260
Db 349 AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal 368
Qy 1261 CAACCTTTGAGATCTTATTGAAGAGTTGAATATATATTGCCAGAGCTGGATGGACTCCT 1320
Db 369 GlnProPheSerSerLeuPheProLysValGluTyrIleAlaArgAlaGlyTrpThrArg 388
Qy 1321 GAGGGAATAATATGCTTGGTCCATCTCTACATGATCCCTCCAGACTCGGCTACAGATAGTG 1380
Db 389 AspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal 408

QY	1381	TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAAGCCAGAGACTC	1440
Db	409	LeuLeuProAlaLeuLeuPheLeuProSerThrGluAenGluGluGluGluAenLeuAse	428
QY	1441	ATTGAGTCAGTCTGATTTCTGTGAGCCCACTAATTTATCTATGAAGAAACCAACAGACATC	1500
Db	429	AlaArgAlaValProArgAenValGlnProTyrValValTyrGluGluValThrAenVal	448
QY	1501	TGGATAAATATCCATGACATCTTTTATCATGTTTTTCCCAAGTCAC---GAAGAGGAAT	1557
Db	449	TrpIleAenValHisAsePheTyrProPheProGlnSerGluGluGluAsePheLeu	468
QY	1558	GAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTCCTGTCATTTATACAAATATACATCT	1617
Db	469	CysPheLeuAenValAenGluCysLeuThrGlyPheCysHisLeuTyrLysValThrAla	488
QY	1618	ATTTTAAAGAAAGCAAAATAAACGATCCAGTGTGGCTGCTCCCTCCCAAGTATTC	1677
Db	489	ValLeuLysSerGlnGlyTyrAseTyrPheSerGluProPheSerProGlyGluAsePhe	508
QY	1678	AGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGTGTGTGATGGAGTTCTTGGCGG	1737
Db	509	LysCysProLeuLysGluGluIleAlaLeuThrSerGlyLysTrpGluValLeuAlaArg	528
QY	1738	CATGGATCTAATATCCAAAGTTGATGAAGTCAGAGCTGGTATATTTTGAAGGCCCAAA	1797
Db	529	HisGlySer-----LysGlyThrLys	535
QY	1798	GACTCCCTTTAGACATCACTGTAGTAGTACGTAAATCCCTGGAGAGGTGACA	1857
Db	536	AspThrProLeuGluHisLeuTyrValValSerTyrGluAlaAlaGlyGluIleVal	555
QY	1858	AGGTGATGACCGTGGCTACTCACATTTCTGTGCATCAGTCAGCAGCTGTGACTCTTT	1917
Db	556	ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAenPheAspMetPhe	575
QY	1918	ATAAGTAAATAGTATGATCAACAGAAATCCACACTGTGTCTCTTTTGAAGTATCAAGT	1977
Db	576	ValSerHisTyrSerSerValSerThrProCysValHisValTyrLysLeuSerGly	595
QY	1978	CCTGAAGATGACCAACTCTCAAAACAAAGAAATTTGGCCACCACTTTGGATTCAGCA	2037
Db	596	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAla	615
QY	2038	GTCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTT	2097
Db	616	SerCysProProAspTyrValProProGluIlePheHisPheHisThrArgSerAspVal	635
QY	2098	ACATTGATGGAGTCTCTACAGCCCTCATGATCTACAGCTCGAAGAAATATCTTACT	2157
Db	636	ArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThr	655
QY	2158	GTGCTGTTCATATATGTTGGTCTCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGAGTC	2217
Db	656	ValLeuPheValTyrGlyGlyProGlnValGlnLeuValAenAenSerPheLysGlyIle	675
QY	2218	AAGTATTTCCGCTTGAATACCTAGCCCTCTCAGGTTATGTGGTGTGTAGTAGACAAAC	2277
Db	676	LysTyrLeuArgLeuAenThrLeuAlaSerLeuGlyTyrAlaValValValIleAspGly	695
QY	2278	AGGGATCTCTGACCCAGGCTTAAATTTGAAGGGCCCTTTAAATATATAAATGGGTCAA	2337
Db	696	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAenGlnMetGlyGln	715
QY	2338	ATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAAT	2397
Db	716	ValGluIleGluAseGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIle	735
QY	2398	GACTTAGATCTGTGGGCATCCCGGCTGTCTTATGGAGGATACCTCTCCCTGTATGGCA	2457
Db	736	AspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGlyPheLeuSerLeuMetGly	755
QY	2458	TTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCAGTCACCTCTGTGG	2517
Db	756	LeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAlaProValThrValTyr	775
QY	2518	ATCTTCTATGATACAGATACAGGACGTTATATGGTCACTCCCTGACCAAGATGAACAG	2577
Db	776	MetalatyrAspThrGlyTyrThrGluArgTyrMetAspValProGluAenAenGlnHis	795
QY	2578	GGCTATTACTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAACCAATCGT	2637
Db	796	GlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuProAenGluProAenArg	815
QY	2638	TTACTGCTCTTACATGGTTCTCTCGATGAGAGATGTCATTTTGCACATACCATATATTA	2697
Db	816	LeuLeuIleLeuHisGlyPheLeuAsePheValHisPheHisThrAenPheLeu	835
QY	2698	CTGAGTCTTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAG	2757
Db	836	ValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAenGluArg	855
QY	2758	CACAGCATAAGAGTTCTCTGAATCGGAGAACATTATGAATTCATCTCTTGTGACTACCTT	2817
Db	856	HisSerIleArgCysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeu	875
QY	2818	CAAGAAAACCTT 2829	
Db	876	GlnGluTyrLeu 879	
RESULT 12			
US-10-070-464-5			
; Sequence S, Application US/10070464			
; Patent No. 6881564			
; GENERAL INFORMATION:			
; APPLICANT: GORRELL, Mark Douglas			
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
; FILE REFERENCE: GH-007			
; CURRENT APPLICATION NUMBER: US/10/070,464			
; PRIOR FILING DATE: 2002-03-07			
; PRIOR APPLICATION NUMBER: PCT/AU00/01085			
; PRIOR FILING DATE: 2000-09-11			
; PRIOR APPLICATION NUMBER: AU PQ5709			
; PRIOR FILING DATE: 2000-02-18			
; PRIOR APPLICATION NUMBER: AU PQ2762			
; PRIOR FILING DATE: 1999-09-10			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; TYPE: PRT			
; ORGANISM: Homo Sapiens			
US-10-070-464-5			
Alignment Scores:			
Pred. No.:	3,528-248	Length:	465
Score:	2422.00	Matches:	465
Percent Similarity:	82.3%	Conservative:	0
Best Local Similarity:	82.3%	Mismatches:	0
Query Match:	43.6%	Indels:	100
DB:	2	Gaps:	1
US-10-825-632-2 (1-3120) x US-10-070-464-5 (1-465)			
QY	1165	ACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATTGCTGAA	1224
Db	1	ThrGlyThrAlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu	20
QY	1225	GGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCAACCTTTGAGATTCTATTGAA	1284
Db	21	GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu	40
QY	1285	GGAGTTGAATATATGTCACAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGGTCCATC	1344

Db 41 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 60
QY 1345 CTACTAGATCGCTCCAGACGCTGCTACAGATAGTGTGATCTCACCTCAATATTATTC 1404
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 80
QY 1405 CCAGTAGAAGATGATGTTATGGAAGGCGAGAGCTCATTTGAGTCAGTCGCTGATCTGTG 1464
Db 81 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100
QY 1465 ACGCCACTAATATTCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 1524
Db 101 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 120
QY 1525 CATGTTTTTCCCAAGTCACGAAGAGAAATGAGTTATTTTCCCTCTGAGTGCAAA 1584
Db 121 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 140
QY 1585 ACAGTTTCCGTCATTTATACAAATTAACATCTATTTTAAAGGAAGCAATATAACGA 1644
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 160
QY 1645 TCCAGTGTGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAGGAGGATAGCA 1704
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180
QY 1705 ATTACAGTGTGTAAGGAGTCTTGGCGCGAGTGAATCAATCAAGTTGATGAA 1764
Db 181 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 200
QY 1765 GTCAGAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAC 1824
Db 201 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 220
QY 1825 GPAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTACCGTGGCTACTCACAT 1884
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240
QY 1885 TCTTCTGCATCAGTCACGACTGTGACTTCTTTAAGTAAGTATAGTAAACCAAGAAGAT 1944
Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 260
QY 1945 CCACACTGTGTCTTACAGCTATCAAGCTCAAGCTCAAGATGACCACTGCAAAACA 2004
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280
QY 2005 AAGGAATTTGGGCCACCATTTTGGATTCAGCAGGCTCTTCTGCTGACTATATCTCTCCA 2064
Db 281 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300
QY 2065 GAAATTTCTCTTTTGAAGTACTGATTTACATTTGATGGATGCTCTACAAGCT 2124
Db 301 GluIlePheSerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320
QY 2125 CATGATCTACAGCTCGAAAGAAATATCTACTGCTGTGTTATATATATGTTGCTCTCAG 2184
Db 321 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 340
QY 2185 GTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGCC 2244
Db 340 ----- 340
QY 2245 TCTCTAGGTTATGTGGTTGTAGTAGATAGACAACAGGGGATCCTGTCAACGAGGGCTTAAA 2304
Db 340 ----- 340
QY 2305 TTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAATTGACGATCAGTTGGAAGCA 2364
Db 340 ----- 340
QY 2365 CTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGC 2424
Db 340 ----- 340

QY 2425 TGTCTCTATGAGGATACCTCTCCCTGATGGCATTAATGTCAGAGGTGAGATATCTTCAGG 2484
Db 340 ----- 340
QY 2485 GTTGTCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAA 2544
Db 341 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 360
QY 2545 CGTTATATGGTCAACCTGAGCAGCAATGAACAGGCTATTACTTAGGATCTGTGGCCATG 2604
Db 361 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 380
QY 2605 CAAGCAGAAAGCTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTCGAT 2664
Db 381 GlnAlaGlyPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 400
QY 2665 GAGAATGTCCATTTTGCACATACCATGATATTTACTGAGTGTATTTTAGTGAGGCTGGAAG 2724
Db 401 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 420
QY 2725 CCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTTGAATCGGA 2784
Db 421 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 440
QY 2785 GAAATATTGAATGCTATCTTTTGCACACTACCTTCAAGAAAACCTTGGATCAGTATTGCT 2844
Db 441 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 460
QY 2845 GCTCTAAAAGTGATA 2859
Db 461 AlaLeuLysValIle 465

RESULT 13

US-09-976-674-29
; Sequence 29, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 29
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-29

Alignment Scores: 2,07e-246 Length: 832
Pred. No.: 2407.00 Matches: 446
Score: 2407.00 Conservative: 125
Percent Similarity: 71.5% Mismatches: 198
Best Local Similarity: 55.8% Indels: 30
Query Match: 43.4% Gaps: 5
DB: 2

US-10-825-632-2 (1-3120) x US-09-976-674-29 (1-832)

QY 124 TGGAGGCGGCGCAGCATGAAGCGGCGCGCTCCATAGCGCAGCTCGGACGCTCC 183
Db 17 TrpArgSerPheSerLeuLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
QY 184 GGGCGGCGCGGGGGAAGGAAAATGCAACATGGCAGCAGCAATGGAACAGAACAGACGCTG 243
Db 37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45

QY	244	GGTGTGAGATATTTGAAATCGCGACTCTGAGGAGAAATATTGAATCACAGGATCGGCCT	303	QY	1321	GAGGAAAAATATGCTTGGTCTCATCTACTAGATCGCTCCAGACTCGCTCAGATAGTG	1380
Db	46	-----ThrAspAspPro	49	Db	389	AspGlyIysIyrAlaIatPalaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal	408
QY	304	AAATTGGAGCCTTTTATGTTGAGCGGTATTCTCGAGTCAGCTTTAAAAAGCTCTCTGCC	363	QY	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Db	50	AlaAlaArg--PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis	68	Db	409	LeuLeuProAlaLeuPheIleProSerThrGluAsnGluGlnGlnArgLeuAlaSer	428
QY	364	GATACCAAGAAATATCATGTGCTACATGATGGCTTAAGGACACCATGATTTTATGTTGTG	423	QY	1441	ATTGAGTCAGTGCTGTGATCTGTGACGCCACTAATTTATCTATGAAGAAAACAAGACATC	1500
Db	69	GlySerArgIysIyrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal	88	Db	429	AlaArgAlaValProArgAsnValGlnProIyrValValIyrGluGluValThrAsnVal	448
QY	424	AAGAGGAATCATCCAGATGACCTCATTCAGACAGAACTATTATCTTGCATCTGCTGGT	483	QY	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCGCCAAAGTCAC---GAAGAGGAAAT	1557
Db	89	GlnIysThrAspGluSerGlyProHisSerHisArgLeuTy-TyrLeuGlyMetProIyr	108	Db	449	TrpIleAsnValHisAspIlePheTyProPheProGlnSerGluGlyGluAspGluLeu	468
QY	484	GAGAACAGAGAAATACACTGTTTATTCTGAATTTCCCAAACTCATCAATAGACAGCA	543	QY	1558	GAGTTTATTTTGGCTCTCGAATGCAAAACAGGTTTTCGTCATTTATACAAAATACATCT	1617
Db	109	GlySerArgGluAsnSerLeuLeuTySerGluIleProIyLeuValArgLysGluAla	128	Db	469	CysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHisLeuTyLysValThrAla	488
QY	544	GTCCTTATGCTCTCTGGAAGCCTCTTTGGATCTTTTCAGCAACTGACTGACTATGGA	603	QY	1618	ATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTGTCCAAAGTATTC	1677
Db	129	LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaIatProHisGly	148	Db	489	ValLeuLysSerGlnGlyTyAspTrpSerGluProPheSerProGlyGluAspGluPhe	508
QY	604	ATGATTCTCGAAGAAGAACTATTAAAGAAAGAAAACGCATTTGGAACTCGGAATT	663	QY	1678	AGTCTCTCTATCAAGAGGAGATAGCAATTACCAGTGGTCAATGGGAAGTTCTTGCCCG	1737
Db	149	ValTySerArgGluGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle	168	Db	509	LysCysProIleLysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArg	528
QY	664	GCTTCTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGAATT	723	QY	1738	CATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAA	1797
Db	169	ThrSerTyAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeu	188	Db	529	HisGlySerLysIleTrpValAsnGluGluThrLysLeuValTyPheGlnGlyThrLys	548
QY	724	TATCACGTAAGAAGTGGAGGCCACAGGATTTTACGCAACAACTTTAAGGCCCAATCTA	783	QY	1798	GACTCCCTTTAGACATCACCCTGATAGTAGTCAAGTTACGTAAATCCTCGAGAGGTGACA	1857
Db	189	PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGlu	208	Db	549	AspThrProLeuGluHisLeuTyValValSerTyValValSerTyGluAlaAlaGlyIleVal	568
QY	784	GTGGAATCTGTTCTCCCAACATACGATGATGATCCAAAATTTATGCCCGCTCATCCAGAC	843	QY	1858	AGGCTGACTGACCGGTGCTACTCATCTTGTGTCATCATGTCAGCTGAGCTGTGACTCTTT	1917
Db	209	IleIysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla	228	Db	569	ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPhe	588
QY	844	TGGATTGCTTTTATACATAGCACGATATTGATATCTAAACATCTGAACACAGAGAAGA	903	QY	1918	ATAAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTCCTTTTCAACAGCTATCAGT	1977
Db	229	PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGlu	248	Db	589	ValSerHisTySerSerValSerThrProCysValHisValTyLysLeuSerGly	608
QY	904	AGGAGACTCATTTATGTGCACATGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCT	963	QY	1978	CCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCA	2037
Db	249	ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAla	268	Db	609	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAla	628
QY	964	GGAGTCGCTACTTGTCTCCAAAGAAATTTGATAGATATTCGCTATGCTGCTGTG	1023	QY	2038	GGTCTCTTCTCTGACTACTCTCCAGAAATTTTCTTTTGAAGAACTACTCTGGATTT	2097
Db	269	GlyValAlaIatPheValIleGlnGluGluPheAspArgPheThrGlyTyTrpIrpCys	288	Db	629	SerCysProProAspTyValProProGluIlePheHisPheHisThrArgSerAspVal	648
QY	1024	CCAAAAGCTGAACAACTCCAGTGGTGGT---AAAATCTTAGAATCTTATATGAAGAA	1080	QY	2098	ACATTTGATGGGATCTCTACAAAGCTCATCATCTACAGCTGCGAAAGAAATATCTCTACT	2157
Db	289	ProThrAlaSerTrpGluGlySerGluGlyLysThrLeuArgIleLeuTyrgLyuGlu	308	Db	649	ArgLeuTyGlyMetIleTyLysProHisAlaLeuGlnProGlyLysLysHisProThr	668
QY	1081	AATGATGAATCTGAGGTGGAATTTATCTGTTTACATCCCTATGTTTGAAGAGGAG	1140	QY	2158	GTGCTGTTTCATATATGCTGCTCCAGGTGAGTGGTGGTGAATAATCGTTTAAAGGAGTC	2217
Db	309	ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluGluArgLys	328	Db	669	ValLeuPheValTyrgLysProGlnValGlnLeuValAsnSerPheLysGlyIle	688
QY	1141	GCAGATTCAATTCGCTTATCTTAAACACAGGTACAGCAAACTCCTAAAGCTCTTTAAGAT	1200	QY	2218	AGTATTTCCGCTTGAATACCTAGCTCTCTAGTTATGTTGTTAGTGTAGTATGACACAAC	2277
Db	329	ThrAspSerTyArgTyProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu	348	Db	689	LysTyLeuArgLeuAsnThrLeuAlaSerLeuGlyTyAlaValValIleAspGly	708
QY	1201	TCAGAAATAATGATTGCTGGAAGGAGGATCATAGTCTCATAGATAAGGAACATAAT	1260	QY	2278	AGGGATCTCTGACCGGCGCTTAAATTGAAGCGCTTTAAATATATAAATCGGTCAA	2337
Db	349	AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal	368	Db	709	ArgGlySerCysGlnArgGlyLeuArgPheGlyAlaLeuLysAsnGlnMetGlyGln	728
QY	1261	CAACTTTTGTAGATCTTATTTGAAGGAGTTGAATATATTCAGAGCTGGATGCTCT	1320	QY	2338	ATAGAAATTCAGCATCAGGTGGAAGACTCCCAATATCTAGTCTCTCCGATATGATTCATT	2397
Db	369	GlnProPheSerSerLeuPheProLysValGluTyIleAlaArgAlaGlyTrpThrArg	388	Db	729	ValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLysTyrgLysPheIle	748
				QY	2398	GACTTAGATCGTGTGGGCATCCACGCGTGGTCTCTATGAGGAGATACCTCTCCCTGATGCCA	2457


```
QY 1024 CCRAAGCTGAACAACACTCCCACTGGTGGT---AAAATTCTTAGAATTCCTATATGAAGAA 1080
Db 289 ProThrAlaSerTrpGluGlySerGluGlyLeuValThrLeuArgGluGluGlu 308
QY 1081 AATGATGAATCTGAGGTGGAATATTATTCATGTTACATCCCTATGTTGGAAACAAGGAGG 1140
Db 309 ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluGluArgLys 328
QY 1141 GCAGATTCAATCCGTTATCCTAAACAGGTACAGCAATCCTAAAGTACCTTTTAAGAT 1200
Db 329 ThrAspSerTrpArgTrpProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu 348
QY 1201 TCAGAAATAATGATTCATGCTGAAGAGGATCATAGATGTCATAGATAGGAACTAATT 1260
Db 349 AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal 368
QY 1261 CAACCTTTGAGATTCATTTGAGAGTTGAATATATGTCAGAGCTGGATGAGTCCCT 1320
Db 369 GlnProPheSerSerLeuPheProLysValGluTrpIleAlaArgAlaGlyTrpThrArg 388
QY 1321 GAGGGAATAATGCTGGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Db 389 AspGlyLysTrpAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal 408
QY 1381 TTGATCTCACCTGAATTTATTATCCAGTAGAAGATGATGTTATGGAAGGACAGACTC 1440
Db 409 LeuLeuProAlaLeuPheIleProSerThrGluAsnGluGluGlnArgLeuAlaSer 428
QY 1441 ATTGAGTCAGTGGTCAATTCGTGACGCCCACTAAATATCTATGAGAAACAACAGACATC 1500
Db 429 AlaArgAlaValProArgAsnValGlnProTrpValValTrpGluValThrAsnVal 448
QY 1501 TGGATAAATATCCATGACATCTTTTCATGTTTCCCAAAAGTCAC---GAAGAGAAATT 1557
Db 449 TrpIleAsnValHisAspIlePheTrpProPheProGlnSerGluGlyGluAspGluLeu 468
QY 1558 GAGTTATATTTTGGCTCTGATGCAAAACAGAGTTCCGTCATTTATACAAAATTACATCT 1617
Db 469 CysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHisLeuTrpLysValThrAla 488
QY 1618 ATTTTAAGAAACCAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAGTGATTTC 1677
Db 489 ValLeuLysSerGlnGlyTrpAspTrpSerGluProPheSerProGlyGluAspGluPhe 508
QY 1678 AAGTGCTCTATCAAGGAGGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGG 1737
Db 509 LysCysProLysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArg 528
QY 1738 CATGGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAA 1797
Db 529 HisGlySer-----LysGlyThrLys 535
QY 1798 GACTCCCTTTTAGACATCACCTGTAGTACGTAGTACGTAAATCTGGAGAGGTGACA 1857
Db 536 AspThrProLeuGluHisLeuTrpValValSerTrpGluAlaAlaGlyGluIleVal 555
QY 1858 AGGCTGACTGACCGTGGTACTCACATTCCTTGTGATCAGTCAGCAGCTGTGACTTCTTT 1917
Db 556 ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPhe 575
QY 1918 ATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGT 1977
Db 576 ValSerHisTrpSerSerValSerThrProProCysValHisValTrpLysLeuSerGly 595
QY 1978 CCTGAAGTACCCCACTTCGAAACAAGGAATTTTGGGCCCAACCATTTTGGATTTCAGCA 2037
Db 596 ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAla 615
QY 2038 GGTCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTT 2097
Db 616 SerCysProProAspTrpValProProGluIlePheHisPheHisThrArgSerAspVal 635
```

Search completed: April 14, 2006, 11:37:46
Job time : 181 secs

```
QY 2098 ACATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTCGAAGAAATATCTCTACT 2157
Db 636 ArgLeuTrpGlyMetIleTrpLysProHisAlaLeuGlnProGlyLysHisProThr 655
QY 2158 GTGCTGTTTATATATGCTGCTCCTCAGTGCAGTTCGGTGAATAATCGGTTTAAAGGAGTC 2217
Db 656 ValLeuPheValTrpGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIle 675
QY 2218 AAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTTGTAGTATGATAGACAAC 2277
Db 676 LysTrpLeuArgLeuAsnThrLeuAlaSerLeuGlyTrpAlaValValIleAspGly 695
QY 2278 AGGGATCCTGTCCACGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAA 2337
Db 696 ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGln 715
QY 2338 ATAGAAATTACGATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCATT 2397
Db 716 ValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLysTrpGlyPheIle 735
QY 2398 GACTTAGATCGTGTGGGCATCCACGCTCGTCTATGAGAGGATACCTCTCCCTGATGCA 2457
Db 736 AspLeuSerArgValAlaIleHisGlyTrpSerTrpGlyGlyPheLeuSerLeuMetGly 755
QY 2458 TTAATGCAGAGGTCAGATATCTTCAGGGTGTCTATTGTGGGGCCCGCAGTCCTCTG 2514
Db 756 LeuIleHisLysProGlnValPheLysAlaGlnProLeuAlaTrpProArgLeu 774
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 14, 2006, 12:23:10 ; Search time 86 Seconds
(without alignments)
3031.692 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5552
Sequence: 1 aagtgctaaagcctccgagg.....agaattactaaaaaaa 3120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+_n2p.model -DEV=xlsp
-Q=/abss/ABSSWEB/spool/US10825632/runat_14042006_092359_3688/app_query.fasta_1
-DB=Published Applications AA Main -QFMT=fastan -SUFFIX=n2p.rapbm
-MTNMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes02p
-USER=US10825632 @CCN 1 1 307 @runat_14042006_092359_3688 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	4700	84.7	882	3	US-09-976-674-1	Sequence 1, Appli
2	4700	84.7	882	4	US-10-054-776-2	Sequence 2, Appli
3	4700	84.7	882	4	US-10-170-789-38	Sequence 38, Appli
4	4700	84.7	882	4	US-10-311-035-9	Sequence 9, Appli
5	4700	84.7	882	4	US-10-072-012-622	Sequence 622, App
6	4700	84.7	882	4	US-10-415-122-6	Sequence 6, Appli
7	4700	84.7	882	4	US-10-825-632-1	Sequence 1, Appli
8	4700	84.7	882	5	US-10-982-512-1	Sequence 1, Appli
9	4528.5	81.6	883	4	US-10-072-012-621	Sequence 621, App
10	3607.5	65.0	690	3	US-09-976-674-7	Sequence 7, Appli
11	3607.5	65.0	690	5	US-10-982-512-7	Sequence 7, Appli

12	3513.5	63.3	661	3	US-09-976-674-11	Sequence 11, Appl
13	3513.5	63.3	661	5	US-10-982-512-11	Sequence 11, Appl
14	3504	63.1	658	3	US-09-976-674-19	Sequence 19, Appl
15	3504	63.1	658	5	US-10-982-512-19	Sequence 19, Appl
16	3236	58.3	613	3	US-09-976-674-21	Sequence 21, Appl
17	3236	58.3	613	5	US-10-982-512-21	Sequence 21, Appl
18	2871	51.7	892	3	US-09-976-674-23	Sequence 23, Appl
19	2871	51.7	892	3	US-09-976-674-27	Sequence 27, Appl
20	2871	51.7	892	5	US-10-982-512-27	Sequence 23, Appl
21	2871	51.7	892	5	US-10-982-512-27	Sequence 27, Appl
22	2871	51.7	892	5	US-10-433-757-12	Sequence 12, Appl
23	2871	51.7	892	4	US-10-415-122-2	Sequence 2, Appli
24	2870	51.7	863	3	US-09-976-674-3	Sequence 3, Appli
25	2870	51.7	863	4	US-10-072-012-619	Sequence 619, App
26	2870	51.7	863	5	US-10-982-512-3	Sequence 3, Appli
27	2862	51.5	863	4	US-10-072-012-224	Sequence 224, App
28	2862	51.5	863	4	US-10-072-012-226	Sequence 226, App
29	2852.5	51.4	580	4	US-10-275-505-2	Sequence 2, Appli
30	2852.5	51.4	580	6	US-11-140-224-2	Sequence 2, Appli
31	2835	51.1	830	4	US-10-415-122-7	Sequence 7, Appli
32	2833	51.0	869	4	US-10-415-122-4	Sequence 4, Appli
33	2821.5	50.8	879	3	US-09-976-674-33	Sequence 33, Appl
34	2821.5	50.8	879	3	US-09-976-674-35	Sequence 35, Appl
35	2821.5	50.8	879	5	US-10-982-512-33	Sequence 33, Appl
36	2821.5	50.8	879	5	US-10-982-512-35	Sequence 35, Appl
37	2422	43.6	465	4	US-10-825-632-5	Sequence 5, Appli
38	2407	43.4	832	3	US-09-976-674-29	Sequence 29, Appl
39	2407	43.4	832	3	US-09-976-674-31	Sequence 31, Appl
40	2407	43.4	832	5	US-10-982-512-29	Sequence 29, Appl
41	2407	43.4	832	5	US-10-982-512-31	Sequence 31, Appl
42	2397	43.2	689	4	US-10-072-012-620	Sequence 620, App
43	2357.5	42.5	819	3	US-09-976-674-37	Sequence 37, Appl
44	2357.5	42.5	819	3	US-09-976-674-39	Sequence 39, Appl
45	2357.5	42.5	819	5	US-10-982-512-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; TYPE: PRT
; LENGTH: 882
; ORGANISM: Homo sapiens
US-09-976-674-1

Alignment Scores:
Pred. No.: 0
Score: 4700.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 84.7%
DB: 3
Length: 882
Matches: 882
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-1 (1-882)

QY 214 ATGCACGACCAATGCAACAGCTGGGTGGATATTTTGAACCTGCGACTGT 273
|||||

Db	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Qy	274	GAGGAGAAATTCGAATCACAGGATCGCCCTAAATTTGGAGCCCTTTTATGTTGACGGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Qy	334	TCCTGGAGTCAGCTTAAAGCTGCTGCCGATACACAGAAATATCATGGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Qy	394	GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProHisSer	80
Qy	454	GACAGAATCTATTACCTTGCCATGCTGTGTGAGAACACAGAGAAATACACTGTTTATCT	513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Qy	514	GAATTTCCCAAACTATCAATAGACGACGAGTCTTAATGCTCTCTTGGAGCCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Qy	574	GATCTTTTTCAGGCACACCTGGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAGA	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Qy	634	GAAGAAGAACCATTTGGAACACAGTCGGAATTCCTTACGATATTATCACCAAGGAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Qy	694	ACATTTCTGTTTCAAGCCGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTAGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCAACATACGGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCAAAATTTATGCCCGCTGATCCAGACTGATGCTGTTTATATACATAGCAACCATAT	873
Db	201	AspProLysLeuCysProAlaAspProAspIleAlaPheIleHisSerAsnAspIle	220
Qy	874	TGGATATCTAATCGTAAACACAGAGAAGAAGAGACTCATTTATGTGCAATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCCAAATGGAAGAGATGCCAGATCAGCTGGAGTGGTACCTTTGTTCTCCCAAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
Qy	994	TTTGATAGATATCTGGCTATTGTTGGTGTCCAAAGCTCAAACTCCCAAGTGGTGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Qy	1054	AAAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113
Db	281	LysIleLeuLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
Qy	1114	ACATCCCTCATGTCGAACCAAGAGGGCGAGATTCATTCCGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAATCTTAAGTCACCTTTAAGATGTCAGAAATATGATGATGCTGAAGGAAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATGTCATAGATAAGNACTAATTCACCTTTTGGATCTCTATTGGAAGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCCAGAGCTGGATGGACTCTCAGGAGAAATATGCTGGTCCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaIleTrpSerIleLeuLeuAsp	380
Qy	1354	CGTCCCCAGACTCCCTTACAGATAGTGTGTGATCTCACTCAATATTATTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	GATGATGTTATGGAAAGCCAGAGACTCATTTGATCGTGCCTGTGATCTGTGAGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATGAAGAAACACACAGACATCTGATATAATATCATGACATCTTTTCATGTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCCAAGCTCACGAAGAGGAAATAGCTTATTTTTCCTCTCAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCAATTTATCAAAATTTACATCTATTATAAGGAAGCAAAATATAAAGCATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GGCTCGCTGCTCCCAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGCAATATACCACT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Qy	1714	GCTGAATGGAAAGTCTTTGGCCGCATCGATCTTAATATCCAAGTTCATGAAGTCACAAG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	520
Qy	1774	CTGGTATATTTGAAGGCCCAAGACTCCCTTTAGAGCATCCCTGTACGTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
Qy	1834	TACGTAATCTCTGGAGAGGTGCAAGGCTGATGACCTGCTGCTACTCACATTTCTCTGCTC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys	560
Qy	1894	ATCAGTTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTGTCCTTTTCAAGCTATCAAGTCTCCTGAAGATCACCAACTGTCACAAACAAAGGAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
Qy	2074	TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTCGGAAGAAATATCTGCTGCTGCTGTTTCATATATGTTGTTGCTCCTCAGGTGCTTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	2194	GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuHisSerLeuGly	680
Qy	2254	TATGCTGTTGTAGTATAGAACACACAGGGGATCCTGTCACCGAGGGCTTAAATTTGAAGC	2313
Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	2314	GCCTTTAAATATATAATGGGTCAAAATPAGAAATTCAGATCAGGTGGGAAGGATCTCAATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	2374	CTAGCTTCTCGATATGATTTTCATTTAGATCTGTTGGGGATCCACGGCTGCTCTAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740

```
QY 2434 GGAGGATACCTCTCCCTGATGACATTAATGACAGAGGTACAGATATCTTCAGGGTGTCTATT 2493
DB 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspPheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTATATG 2553
DB 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCAACCTGACAGAGTGAACAGGGCTATTACTAGATCTGTGGCCATGCAAGCAGAA 2613
DB 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTCTCGGATGAGATGTC 2673
DB 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACAGATATATTACTGAGTTTTTTAGTGGGGCTGGAAAGCCATATGAT 2733
DB 821 HisPheAlaHisThrSerLeuLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCCTCAGGAGACACAGACATAAGAGTTCTGAACTCGGAGAACATTAT 2793
DB 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACTGCATCTTTTGCACCTTCAAGAAACCTTGGATCACGTATTTGCTCTCAAAA 2853
DB 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
DB 881 ValIle 882
```

RESULT 2

```
US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2
```

```
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 4 Gaps: 0
```

US-10-825-632-2 (1-3120) x US-10-054-776-2 (1-882)

```
QY 214 ATGGCAGCAGCAATGAAACAGACAGCTGGGTGTGAGATATTTGAACTGGCGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATCACAGGATCGGCTAAATTGGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAAGCTTAAAGCTGCTTGGCCGATACCAGAAATATCATGCTACATCATG 393
DB 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
```

```
QY 394 GCTAAGCCACCATGATGATTTTCATGTTTCTGAAGAGGAATGATCCAGATGGACCTCATTC 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTTATTACTTCCCATGTCGTGAGAACAGAGAAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTCGAAGCCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGGAATGTATTTCTCGAGAGAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuLeuArg 140
QY 634 GAAAGAAAACGCATTTGGAAACAGTCGGAATTTGCTTACGATTATCACCAAGGAAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATACGTAAAGATGAGAGGCCACAAAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTACGCAACAACTTTAAGCCCAATCTAGTGGAAACTAGTTGTGCCAATACCGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATATGCCCCCTGATCCAGACTGGATGCTTTTATATACATAGCAACGATTT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCTGTAACACAGAGAGAGAGACTCACTTATGTGCAATGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAAACATGGAAGAAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGTAGATATTTCTGGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGATTTCTATATGAAGAAATGATGATCTGAGGTGGAAATTTATTCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTTATGTTGAAAACAGAGGGCAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAACTCTTAAGTCACTTTTAAGATGTGAGAAATAATGATGCTGTAAGGAGGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGCTCATAGATAAGGAACCTAAATTCAACCTTTTGAGATTTCTATTTGAAGGAGTTCAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGGAATCTCTGAGGGAAAATATGCTGTGCTCCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAAGGAGAGACTCATTTAGTCCAGTCCCTGATTTCTGTGACCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATCAAGAAACACAGACATCTCGATAAATATCCATGACATCTTTTCATGTTT 1533
```


;
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-170-789-38

Alignment Scores:

Pred. No.:	0	Length:	882
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	84.7%	Indels:	0
DB:	4	Gaps:	0

US-10-825-632-2 (1-3120) x US-10-170-789-38 (1-882)

QY	214	ATGCGCAGCAATGCAACAGAACAGCTGGGTGTGAGATATTGAAACTGCGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGAAATTAATCACAGGATCGCCTAAATGGAGCCCTTTTATGTGGCGGTAT	333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGTTAAAGCTGTTCGCCGATACAGAAATATCATGGCTACATGATG	393
DB	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGCCACCATGATTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGATCTATTACCTTCGATGCTGCTGGAGAACAGAGAAATACACTCTTTTATTCT	513
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAAATCCCAAACTCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCCTCTTTTG	573
DB	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAAAGACTTAAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg	140
QY	634	GAAGAAACCGCATTTGAACAGTCGGAATGCTTCTTACGATTATCACCAAGGAGTGGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTCTGTTTCAAGCCGCTAGTGGAATTTTATCAGCTAAAGATGGAGGCCCAAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTGCCCAACATACGGATG	813

DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAACATCGTAAACCAGAGAGAAAGAGACTCACCTTATGTGCACAATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluArgGluLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAAATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
QY	994	TTTGATAGATTTCTGGCTATTGGTGTCTCCAAAGCTGAAACACTCCAGAGTGGTGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATTCATGTT	1113
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAACAGAGGCGAGATTTCATTCCTCGTTATCTTAAACAGGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCCTAAAGTCATTTTAAGATGTGAGAAATAATGATGATGCTGAAGGAAGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATCTCATAGATAAGGAACATAATCAACCTTTTGAGATTCATATTTGAAGGATGAA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGCTGGATGGACTCTGAGGGGAAATATGCTTGTGCTCATCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCAGTAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATCATGTTATGGAAGGAGAGACTCATTTGAGTCTGAGTCCCTGATTTCTGTGACGCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCATGACATCTTTTCATGTTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAGTCACGAAGAGGAAATGAGTTTATTTTTCCTCTGATGAAATGAACAGGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCAATTATACAAAATTACATCTATNTTTAAGGAAGCAAAATATAAAGATCCAGTGGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGCTGCTGCTCCAGTGTATTTCAAGTGTCTCTATCCAAAGAGAGATGACATTTACCACT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GSTCAATGGGAAGTTCTTGGCGCATGGATCTAATATCAAGTTGTGATGAAGTCAGAAAGG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTGTAAGGCACCAAGACTCCCTCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
DB	521	LeuValTyrPheGluGluThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
QY	1834	TACGTAAATCTCGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560

Qy 1894 ATCGTCAGCACTGCTGACTTCTTTATAGTTRAGTATAGTAACACAGAAAGATCCACACTGT 1953
Db 561 IleserGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy 1954 GTGTCCCTTTACAAGCTATCAAGTCTGAAGATGACCCAACTTCGMAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerProGluAspPheProThrCysLysThrLysGluPhe 600
Qy 2014 TGGGCCACCAATTTGGATTGACAGAGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Qy 2074 TCTTTTGAAGTACTACTCGAATTTACATTTAGGATGCTCTCAAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Qy 2134 CAGCCTCGAAGAAATATCTACTGTCTGTTTCATATATGGTGGTCTCAGGTGAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Qy 2194 GTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Qy 2254 TATGTGGTGTAGTAGTAGACAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Qy 2314 GCCTTAAATATAAATGGGTCAAAATAGAAATTCAGCATCAGTGAAGGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Qy 2374 CTAGCTTCTCATATGATTTCACTTAGCTAGATCGTGTGGGATCCAGCGTGGTCCAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Qy 2434 GGAGGATACCTCTCCGATGGCATTAATGCAGAGTTCAGATATCTTCAGGTTCTCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Qy 2494 GCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAAGTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyThrGluArgTyrMet 780
Qy 2554 GGTCAACCTTGACAGAATGAACAGGGCTATTACTAGGATCTGTGGCCATGCACAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Qy 2614 AAGTTCCTCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Qy 2674 CATTTTGCATACACAGTATATTTACTGAGTTTTTTTGTAGGGCTGGAAGCCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Qy 2734 TTACAGATCTATCTCCAGGAGACACAGCATAAAGTTCCTGAATCGGAGAACATTTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Qy 2794 GAAGTCATCTTTTGCACCTTCAAGAAACCTTGATCAGCTATTTGCTGCTCTTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Qy 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 4
US-10-311-035-9
; Sequence 9, Application US/10311035
; Publication No. US20040023243A1

GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Aneena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannie R.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuning
; APPLICANT: REDDI, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 ECT
; CURRENT APPLICATION NUMBER: US/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CDI
US-10-311-035-9
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 4 Gaps: 0
US-10-825-632-2 (1-3120) x US-10-311-035-9 (1-882)
Qy 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGTGTGTGAGATATTTGAAACTGCGGACTGT 273
Db 1 MetAlaAlaIleMetGluThrGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAAATATTGAATTCACAGATCGGCTAAATTTGGAGCCTTTTATTTTGTGACCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Qy 334 TCTTGAGTCAGCTTAAAGCTGCTTCGCCGATACCGAAATATCATGCTACATCATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGCCACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATCGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy 454 GACAGATCTATTACCTTCCCATGTCTGCTGAGAACAGAGAAATACATCTTTTATCT 513

Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAenThrLeuPheTyrSer	100
Qy	514	GRAATCCCAAAATCATATAGACAGCAGTCTTAATGCTCTCTGGAAGCCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Qy	574	GATCTTTTTCAGGCAACACTCGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAG	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
Qy	634	GAAGAAAACCGATTGGAACAGTCGGAAATTGCTTTACGATTATCACCAAGGAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Qy	694	ACATTTCTGTTCAAGCCGGTAGTGGAAATTATATACGTAAAGATGAGAGGCCACAAAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCACATACCGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCCCAAAATATGCCCCCTGATCCAGACTGGATTGCTTTATACATAGCAACATATT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Qy	874	TGATATCTAACTCTTAACAGAGAGAAAGAGACTCACTATGTGCAATAGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCCAACTGAAGAAGATCCAGATCAGCTGAGAGTCGCTACCTTTGTTCTCCAAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Qy	994	TTTCATAGATATTCTGGCTATTGGTGTGTCAAAGCTGAAACAACTCCAGTGGTGGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Qy	1054	AAATTTCTTAGAATCTATATGAGAAATGATGATCTGAGGTGGAAATTTATTCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
Qy	1114	ACATCCCTATGTTGGAACAGGAGGGCAGATTCAATTCCTGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAATCTTAAGCTACTTTTAAGATGTCAGAAATAATGATTGATCTGGAAGGAAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATGTCATAGATAAGAACTAAATTCAACTTTTGAGATTTTATTTGAAGGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCCAGAGCTGGAGTCTCCTGAGGAAATAATGCTGGTCCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
Qy	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	GATGATGTTATGAAGGCAGAGACTCATTGATGAGTGCCTGATTCTGTGACGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATCAAGAAACACAGACATCTCGATAAATATCCATGACATCTTTTCATGTTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCCAAAGTCACCAAGAGGAATTGAGTTTATTTTGTGCTCTGAAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAAAATATAAACCAGTCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAAAGAGGAGATGCAATTAACAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Qy	1714	GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTATATATCCAAGTTGATGAAGTCAGAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	1774	CTGCTATATTTTCAAGGCCCAAGACTCCCCTTTAGAGACATCACTCTGCTAGCTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
Qy	1834	TACGTAATCCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACTTTCTTGTCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	1894	ATCAGTCAGCACTGTGACTTCTTTATAAGTATAGTAAACCAGAGAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTGTCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCCACTTGCAGAAAACAAAGAAATT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCTGACTATATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Qy	2074	TCCTTTCAAAAGTACTCTGATTTACATTTGATGATGATGCTCTACAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTCGAAAGAAATATCTACTGTGCTGTTTCAATATGTTGGTGGTCTCAGGTGCGAGTTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATFACCTAGCCTCTTAGT	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Qy	2254	TATGTGTTCTAGTGATAGACAACAGGGATCTCTGTCCACCGAGGCTTAAATTTGAAGGC	2313
Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	2314	GCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGTTCGGAAGGACTCCAATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	2374	CTAGCTTCTCGATATGATTTTCAATTGACTTTAGATCGTGGGCATCCAGGCTGGTCCCTAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
Qy	2434	GGAGGATACCTCTCCCTGATGGCATTTAATGACAGAGGTACAGATATCTTCAGGCTTCTATT	2493
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Qy	2494	GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATG	2553
Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Qy	2554	GGTCACCCCTCACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCATGCAACAGACAA	2613
Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Qy	2614	AAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAAATGTC	2673
Db	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820

QY 2674 CATTTCACATACCATATATCTAGTCTTTTGTAGTGGCTGGAAGCCATATGAT 2733
Db 821 HSPheAlaHisThrSerLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGACACACAGATTAAGATTCCTGAATCGGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACCTGATCTTTTGCATCTACCTTCAAGAAACCTTGGATCACGTATTGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 5

US-10-072-012-622 Application US/10072012

Publication No. US20040033493A1

GENERAL INFORMATION:

APPLICANT: Tchernev, Velizar

APPLICANT: Spytek, Kimberly

APPLICANT: Zernusen, Bryan

APPLICANT: Patturajan, Meera

APPLICANT: Shinkets, Richard

APPLICANT: Li, Li

APPLICANT: Gangolli, Bsha

APPLICANT: Padigaru, Muralidhara

APPLICANT: Anderson, David W.

APPLICANT: Rastelli, Luca

APPLICANT: Miller, Charles E.

APPLICANT: Gerlach, Valerie

APPLICANT: Taupier Jr, Raymond J.

APPLICANT: Gusev, Vladimir Y.

APPLICANT: Colman, Steven D.

APPLICANT: Wolenc, Adam R.

APPLICANT: Pena, Carol E. A

APPLICANT: Furtak, Katarzyna

APPLICANT: Grose, William M.

APPLICANT: Alsobrook II, John P.

APPLICANT: Lepley, Denise M.

APPLICANT: Rieger, Daniel K.

APPLICANT: Burgess, Catherine E.

TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

FILE REFERENCE: 21402-258

CURRENT APPLICATION NUMBER: US/10/072,012

CURRENT FILING DATE: 2002-01-31

PRIOR APPLICATION NUMBER: 60/265,102

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 60/265,514

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,517

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,412

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,395

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/266,406

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 60/266,767

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/267,057

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/266,975

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/267,459

PRIOR FILING DATE: 2001-02-08

TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-622
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-072-012-622 (1-882)

QY 214 ATGCACGACCAATGGAACACAGAACAGCTGGGTGTGGAGATATTTGAACTCCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATGAATCACAGGATCGGCTAAATTTGAGCCCTTTTATGTGTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGGATACCCAGAAAATATCATGGCTACATGATG 393
Db 41 SerTyrSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCCACATGATTTTCATGTTTTTGTGAAGAGAAATGATCCAGATGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGGCATGCTGTGTGAGAACAGAGAAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAACTATCAATAGACAGCAGTCTTAAATGCTCTCTTGGAGGCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCAGAGAAGAACTATTAAAG 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAGAAACAGCATTTGGAACAGTTCGGAATTCCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTCTTCAAGCCGCTAGTGGAAATTTTATCACGTAAAGATGGAGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTATCCGCAACACTTTAAGCCCACTAGTGGAACTAGTGGAACTAGTGGAACTAGTGGAA 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCCGCTGATCCAGACTGGATTCGTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCGTAAACAGAGAAGAGAGACTCATTATGTGCACATGACCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAAACATGGAAGAAGATGCCAGATCAGCTGAGTTCGCTACTTTGTTCTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTTCTGGCTATTGTTGGTGTCCAAAGCTGCAAAACCTCCAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300

QY	1114	ACATCCCTATGTTGGAACACAGAGGGGAGAGATTCATTCCGTTTCCCTAAACACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGGAAGGAGGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCTATTGAAAGGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGAGCTCTCTGAGGGAATAATGCTTGCTGCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGATTTATTTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATCATCTTATGGAAGGACAGACTCATTGAGTCAGTGCCTGATTCTGTGAGCGCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACACAGACATCTGATAAATATCATGACATCTTTTCATGTTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCACGAAGGAAATTGAGTTTATTTTTGCTCTGAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAsaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTATACAAATTAATCATCTATTTAAAGGAAGCAATATAACGATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSergly	480
QY	1654	GGGTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAGAGAGAGATACCAATACCAAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAAGTCTTGGCCGCGCATGGATCTAATCAAGTTGATGAAGTCAGAAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTACTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
QY	1834	TACGTAAATCCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAAACAGAAAGATCCACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAGAAACAGGAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACATTTTGGATTACAGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspLysThrProGluIlePhe	620
QY	2074	TCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAGGCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCCTGGAAAGAAATATCTTACTGTGCTGTTCATATATATGGTGTCTCAGGTGCGATTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660

QY	2194	GTGATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGT	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	2254	TATGTGGTTGTAGTATAGACAAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTGAAGC	2313
Db	681	TyrValValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly	700
QY	2314	GCCTTTAAATATAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	2374	CTAGCTCTCCATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
QY	2434	CGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTTCAGGGTGTCTATT	2493
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
QY	2494	GCTGGGGCCCCAGTCCTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATG	2553
Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
QY	2554	GGTCACCTGACCAGAAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA	2613
Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
QY	2614	AGTTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAATGTC	2673
Db	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
QY	2674	CATTTTGACATACACAGTATATATTACTGAGTTTTTTAGTGGGCTGGAAAGCCCATATGAT	2733
Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
QY	2734	TTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGGAGAACATATAT	2793
Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
QY	2794	GAACTGATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAA	2853
Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
QY	2854	GTGATA 2859	
Db	881	ValIle 882	
RESULT 6			
US-10-415-122-6			
; Sequence 6, Application US/10415122			
; Publication No. US20040053369A1			
; GENERAL INFORMATION:			
; APPLICANT: THE UNIVERSITY OF SYDNEY			
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
; FILE REFERENCE: FP15217			
; CURRENT APPLICATION NUMBER: US/10/415,122			
; CURRENT FILING DATE: 2003-08-07			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 6			
; LENGTH: 882			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-415-122-6			
Alignment Scores: -			
Pred. No.:	0	Length:	882
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	84.7%	Indels:	0
DB:	4	Gaps:	0

US-10-825-632-2 (1-3120) x US-10-415-122-6 (1-882)

QY 214 ATGGCAGCAGCAATGGAAACAGACACAGCTGGGTGGTGTGAGATATTGAACTGCGCACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATTGAAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspAArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACAGAAAAATATCATCGCTTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTTCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACCTTGCATGTCTGGTGTGAGAACAGAGAAAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAACTATCAATAGACGACGAGCTTAAATGCTCTCTTGGAGGCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACATGGACTATGGAAATGATTTCTTCGAGAGAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAACGCATTGGAAACAGTCGGAATTCCTTACGATTTATCCAGAAAGAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGGAGGGCCACAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813
Db 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysPheAsnIleArgMet 200
QY 814 GATCCAAAATTTAGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACAGAGAAAGAGACTCACCTTATGTGCACAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGCTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATCTATATGAAGAAATGATGATCTGAGCTGGAATATTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGGAAAACAAGAGGCGAGATTCAATCCGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCTAAAGTCACTTTTAAAGATGTCAGAAATATGATTCATGCTGAAAGGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGTAAGGAACCTAAATTCACACCTTTTGAGATTCTATTGAAAGGAGTTGAA 1293

Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGACTCCTGAGGAAATATCTTGCTCCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAAGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGAGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCAGAAAGAGAAATTTGAGTTTATTTTTTGCCTCTGGAATCAAAACAGAGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCAATTTATACAAATTTACATCTATTATTTAAAGGAAAGCAAAATATAACGATCCAGTGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTCGCTCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGAGGTCTTGGCGCATGGATCTAATATCCAACTGTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCCCAAAAGACTCCCTTTTAGAGCATCCCTGTACGTAGTCACT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAAACCAAGAAATCCCACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTCAAGCTATCAAGTCCTGAAGATGACCAACTTGCACAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTGACTATATCTCTCAGAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTTGTATGGATGCTCTCAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGGAAAGAAATATCTACTGCTGCTGTTCATATATGTTGCTCCTCAGGTGCAAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATATCGGTTTAAAGGAGTCAGTATTTTCCGTTTGAATACCTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTTGTAGTAGATAGACAACAGGGGATCTGTGTCCAGAGGGCTTTAAATTTGAAGGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATCGGTCAAATAGAAATTCACGATCAGGTGGAGGACTCCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720

QY 2374 CTAGCTTCTCGATATGATTTTCATTGACTTGGTGGGATCCACGGCTGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyIlePheTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCAATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GTTCACTCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGACAGAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AGTTCCCTCTGAACCAAACTGTTACTGCTCTTACATGGTTTCTGGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCATACACAGATATATCTGAGTTTTTTTACTGAGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCTCAGGAGACACAGACATAAGAGTTCTCTGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACCTGCATCTTTTGCATCACTTCAAGAAAACCTTGGATCACGTATTCGTCTPAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 7

US-10-825-632-1
; Sequence 1, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-1

Alignment Scores:

Pred. No.:	0	Length:	882
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	84.7%	Indels:	0
DB:	4	Gaps:	0

US-10-825-632-2 (1-3120) x US-10-825-632-1 (1-882)

QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAAACTGCCGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTTGAATTCACAGGATCGGCCTAAATTTGGAGCCCTTTTATTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCTGTGAGTCAAGTTAAAAGCTGCTGCCGATACCCAGAAAATATCATCGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCCACCATGATTTTCATGTTTGTGAAGGATGATCCAGATGGACCTCATTTCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAAATCTATTACCTTCCCATGTCTGTGAGAACAGAGAAAATACACTGTTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCCAAAATATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGGACTATCGAATGTATTCTCGAAGAGAAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCATTTGAAACAGTCGGAATTCCTTTACGATTATACCAAGGAAGTGGGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGTTAGTGAATTTATCAGTAAAGATGGAGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATPAGCAACGATTT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAAACAGAGAAAGAGAGACTCATTATGTGCACAACTAGACTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCGTGCTATTGCTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGCGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTTATGTTGGAAAACAGGAGGCGAGATTCATTCCCGTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAAATAATCATGTGATGCTGAAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCTATAGATAAGGAACCTAATTCACCTTTTGAGATTCATTATTTGAAGAGGTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360

QY	1294	TATATTGCCAGAGCTGGATGGACTCCTGAGCGGAAAATATGCTTGGTCCACTCTACTAGAT	1353	Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	QY	2434	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTCGTATT	2493
QY	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTTATCCCAAGTAGAA	1413	Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	QY	2494	GCTGGGGCCCGCAGTCATCTCTGGATCTTCTATGATACAGGATACACGGAAAGCTTATATG	2553
QY	1414	GATGATCTTATGAAAGCAGAGACTCATTGAGTCAGTCCCTGATTCCTGTGACGCCACTA	1473	Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	QY	2554	GGTCACCTCTGACCAAGATGAACAGGGCTATTACTTATAGGATCTGTGGCCATGCACAGCAAA	2613
QY	1474	ATTATCTATGAAGAAACACAGACTCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533	Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	421	IleIleTyrGluGlnThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	QY	2614	AAGTTCCCTCTGAAACCAAAATCGTTTACTCTCTTACATCGTTTCTCGATGAGAATGTC	2673
QY	1534	CCCCAAAGTCACGAAGAGAAATTGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTC	1593	Db	801	LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	QY	2674	CATTTTGCACATACACAGTATATTACTGAGTTTTTTTAGTAGGGCTGGAAGCCATATGAT	2733
QY	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGT	1653	Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	QY	2734	TTACAGATCTATCCTCAGGAGAGACACAGCATAAAGAGTTCTCTGAATCGGAGAACATTAT	2793
QY	1654	GGCTGCTCTGCCAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGCAATATACCAAGT	1713	Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	QY	2794	GAAGTCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAAA	2853
QY	1714	GGTGAATGGGAAGTTCTTGGCCGCATGGATCTTAATATCCAAGTTGATGAAGTCAGAAGG	1773	Db	861	GluLeuHisLeuLeuHisTyrIleGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	QY	2854	GTGATA 2859	
QY	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833	Db	881	ValIle 882	
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540	RESULT 8			
QY	1834	TACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893	US-10-982-512-1			
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	; Sequence 1, Application US/10982512			
QY	1894	ATCAGTCAGCTGTGACTCTTTTATAAGTAAGTATAGTAACTAACCAAGAAGAAATCCACACTGT	1953	; Publication No. US2005005081A1			
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	; GENERAL INFORMATION:			
QY	1954	GTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGGAATTT	2013	; APPLICANT: Qi, Steve			
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	; APPLICANT: Akinsanya, Karen			
QY	2014	TGGGCCACCATTTTGGATTCACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC	2073	; APPLICANT: Riviere, Pierre			
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	; APPLICANT: Junien, Jean-Louis			
QY	2074	TCTTTTGAAGTACTACTCGATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTA	2133	; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	; FILE REFERENCE: 70669			
QY	2134	CAGCTCGGAAAGAAATATCCACTGCTGTCTTATATATGCTGTCTCAGGTGCGAGTTG	2193	; CURRENT APPLICATION NUMBER: US/10/982,512			
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660	; CURRENT FILING DATE: 2004-11-05			
QY	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGT	2253	; PRIOR APPLICATION NUMBER: US/09/976,674			
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680	; PRIOR FILING DATE: 2001-10-12			
QY	2254	TATGTGGTTGTAGTGATAGACAAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGGC	2313	; PRIOR APPLICATION NUMBER: US 60/240,117			
Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700	; PRIOR FILING DATE: 2000-10-12			
QY	2314	GCCTTTAAATATAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAGGACTCCATAT	2373	; NUMBER OF SEQ ID NOS: 61			
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720	; SOFTWARE: PatentIn version 3.1			
QY	2374	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGGTGGGCATCCACGGCTGGTCTAT	2433	; SEQ ID NO 1			
				; LENGTH: 882			
				; TYPE: PRT			
				; ORGANISM: Homo sapiens			
				US-10-982-512-1			
				Alignment Scores:			
				Pred. No.:		0	
				Score:		4700.00	
				Percent Similarity:		100.0%	
				Best Local Similarity:		100.0%	
				Query Match:		84.7%	
				DB:		5	
						Length: 882	
						Matches: 882	
						Conservative: 0	
						Mismatches: 0	
						Indels: 0	
						Gaps: 0	
				US-10-825-632-2 (1-3120) x US-10-982-512-1 (1-882)			
QY	214	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGTGAGATATTTGAACTCGGACTGT	273				
Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20				

QY	274	GAGGAGAAATTTGAATTCACAGGATCGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProIysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAGAGCTGCTCCGATACCGAAGAAATATCATGCTCATCATGATG	393
Db	41	SerTrpSerGlnLeuIysIysLeuLeuAlaAspThrArgIysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA	453
Db	61	AlaIysAlaProHisAspPheMetPheValIysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAACTATTACCTGCCATGCTCGTGGAGAACACAGAGAAATACACATGTTTATTCT	513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAATTTCCCAAACTATCAATAGACGACGACTTAAATGCTCTCTGGAGGCTCTTTTG	573
Db	101	GluIleProIysThrIleAsnArgAlaAlaValleuMetLeuSerTrpIysProLeuLeu	120
QY	574	GATCTTTTTCAGGCAACATCGGACTATGGAATGATTCTCGAGAGAAAGAACTATTAAAG	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
QY	634	GAAGAAACGCATTTGGAACAGTCGGAATTCCTTACGATTATCACCAAGGAAGTGA	693
Db	141	GluArgIysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGAGTAGTGAATTTATCACGTAAGAGATGAGGGCCCAAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValIysAspGlyGlyProGlnGly	180
QY	754	TTTACGCCAACACCTTTAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACCGATG	813
Db	181	PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATTTATGCCGCTGATCCAGACTGGATTCCTTTATACATAGACACGATATT	873
Db	201	AspProIysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAACATCGTAACACAGAGAAGAGAGACTCACTTATGTCACAAATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValleuGlnGluGlu	260
QY	994	TTTGTAGATATTCTGGCTATTGGTGGTGTCTCAAAAGCTGAAACAACTCCCAAGTGGTGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProIysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAAATCTTGAATTTCTATATGAAGAAATCATGAATCTGAGGTGGAAATTTATTCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAACAAGGAGGCGCATTCATTCCGTTATCTCTAAACAGAGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProIysThrGlyThr	320
QY	1174	GCNAATCTTAAAGTCACTTTTAAGATGTCAGAAATATGATGCTGAGGAGGAGATC	1233
Db	321	AlaAsnProIysValThrPheIysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGGAACATAATCAACCTTTGAGATTCCTATTGGAAGAGTTGAA	1293
Db	341	IleAspValIleAspIysGluLeuIleGlnProPheGluIleLeuPheGluValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTGGTCCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyIysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCCAAGATCGCCTACAGATAGTGTGATCTCACCTGAATATTATTTATCCCAAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValleuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGGAAGGCGAGACATCATTTAGTCAGTCGCTGATTCTGTGACGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProIleu	420
QY	1474	ATTATCTATGGAAGAAACACACATCTCGGATAAATATCCATGACATCTTTTCATGTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAGTCACCAAGAGGAAATTCAGTTTATTTTGGCTCTGTAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysIysThrGlyPhe	460
QY	1594	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAATCCAGTCCAGTGT	1653
Db	461	ArgHisLeuTyrIysIleThrSerIleLeuIysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGGCTGCCTCTCAAAGTCATTTCAAGTGTCTCTCAAGAGGAGATAGCAATACCAGT	1713
Db	481	GlyLeuProAlaProSerAspPheIysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAAGTCTTCGCGGCGATGATCTAATATCTCAAGTTGATGAAGTCAGAAG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuLeuTyrValValSer	540
QY	1834	TAGTAAATCTCGAGAGGTGACAGGCTGACCTGACCTGCTACTCATCTTCTGTGTC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTCACTCTTTTATAGTAAGTATAGTACCAAGAGAAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCCTTTTACAAAGCTATCAAGTCCTCAAGATGCCCAACTTGCACAAACAAAGAAATTT	2013
Db	581	ValSerLeuTyrIysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCCAACCATTTTGGATTCAGCAGGTCTCTTCTCTGATCTATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	2074	TCCTTTGAAAGTACTACTGGATTACATGTTATGGATGGCTCTACAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTCGGAAGAAATATCTACTGCTGTTTCATATATATGTTGCTCCTCAGGTGCACATG	2193
Db	641	GlnProGlyIysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	2194	GTGAATAATCGGTTTAAAGAGGTCAAGTATTTCCCTTGAATACCTAGACCTCTCTAGGT	2253
Db	661	ValAsnAsnArgPheIysGlyValIysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	2254	TATGTGGTGTCTAGTATAGACCAACAGGGGATCTCTGTCCAGCGGGCTTTAAATTTGAAGGC	2313
Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
QY	2314	GCCTTTAAATATATAATGGGTCAAAATAGAAATTCAGATCAGGTGGAGGACTCCAAATAT	2373
Db	701	AlaPheIysTyrIysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	2374	CTAGCTTCGATATGATTTTCATTTAGTATCTGTGGGATCCAGCGGTGGTCTCTAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
QY	2434	GGAGGATACCTCTCCCTGATGGCATTATATGCAGAGGTCAATATCTTCTCAGGTTCTATT	2493

Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Qy 2494 GCTGGGCCCCAGCTCTCTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Qy 2554 GGTCCACCTACAGCATGACAGGCTATTACTTATAGGATCTGTGGCCATCAAGCAGAA 2613
Db 781 GlyHisProAspGlnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Qy 2614 AAGTCCCTCTGACCAATCTTTACTGCTCTACATGTTTCTCGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Qy 2674 CATTTTGCACATACAGTATATTAATCTAGTCTTTTATGAGGGCTGGAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Qy 2734 TTACAGATCTATCTCTAGGAGACACAGCATAAAGATTCTCGAATCGGGAGACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Qy 2794 GAATGTCATCTTTTGCACCTACCTTCAAGAAAACCTTTGGATCAGTATTGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Qy 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 9
US-10-072-012-621
; Sequence 621, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Ichernev, Velizar
; APPLICANT: Spycek, Kimberly
; APPLICANT: Zerkhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinketsa, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072, 012
; PRIORITY FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 621
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-621
Alignment Scores:
Prod. No.: 0 Length: 883
Score: 4528.50 Matches: 845
Percent Similarity: 97.7% Conservative: 18
Best Local Similarity: 95.7% Mismatches: 19
Query Match: 81.6% Indels: 1
DB: 4 Gaps: 1
US-10-825-632-2 (1-3120) x US-10-072-012-621 (1-883)
Qy 214 ATGGCAGCAGCATGGAACAGACAGACAGCTGGGTGTGAGATATTTGAAACTCGCGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGlnMetGlyValGluIlePheGluThrAlaGluCys 20
Qy 274 GAGGAG--AATATTGAATCAACAGATCGCGCTTAATTTGGAGCTTTTATTTGTTAGCGG 330
Db 21 GluGluGlyAsnGlyGluSerGlnAspArgProLysLeuGluProPheTyrValGluArg 40
Qy 331 TATTCCTGAGTCAGCTTTAAAAAGCTGCTTCCGATACACAGAAAATATCATGCTACATG 390
Db 41 TyrSerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMet 60
Qy 391 ATGGCTAAGGCACCATGATTTTCATGTTGTGAAGAGGAATGATCAGATGACCTCAT 450
Db 61 MetAlaLysAlaProHisAspPheMetPheValLysArgThrAspProAspArgProHis 80
Qy 451 TCAGACAGAAATCTATTACCTTTGCCATGTCTGTCGAGAACACAGAGAAAATACACTGTTTAT 510
Db 81 SerAspArgValTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 100
Qy 511 TCTGAAATTCACAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCTCTT 570
Db 101 SerGluLeuProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeu 120
Qy 571 TTGGATCTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAACTATTATTA 630
Db 121 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeu 140
Qy 631 AGAGAAGAAACGATGGAACAGTGGACAGTCCGAATGCTTCTTACGATTATCACCAAGGAAGT 690
Db 141 ArgGluArgLysArgIleGlyThrValGlyIleAlaAlaTyrAspTyrHisProGlySer 160
Qy 691 GGAAACATTTCTTTCAACCCGCTAGTGAATTTATACGTTAAAGATGGAAGGCCCAAA 750
Db 161 GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisIleLysAspGlyProHis 180
Qy 751 GGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTGTCCCAACATACGG 810
Db 181 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 200
Qy 811 ATGGATCCAAAATTTATGCCCGCTGATCCAGACTCGATTGCTTTTATATCATAGCAAGAT 870
Db 201 MetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAsp 220

Qy	871	ATTTCGATATCTAACATCGTAACAGAGAAGAAAGAGACTCACTTATGTGCCAANTGAG	930
Dd	221	IleTpIleSerAsnLeuValThrArgGluGuaArgIleThrTyValHisAsnGlu	240
Qy	931	CTAGCCAACTGAAGAAGATGCCAGNATCAGCTCAGCTGGAGTCGCTACCCTTGTTCTCCAAGAA	990
Dd	241	LeuAlaAsnMetGluGluAppProArgSerAlaGlyValAlaThrPheValLeuInGlu	260
Qy	991	GAAATTTGATAGATATTCTTGCGCTATTGTGTGTGTCTCCAAAAGCTGAAAAACAACCTCCCAGTGGT	1050
Dd	261	GluPheAspAa-gtyrSerGlyTyTrtpCyseProGlnalaGluArgThrProSerGly	280
Qy	1051	GGTAAAAATCTTAGAAATTTCTATATAGAAAAATGATGAATCTGAGGTGGAAATTTATTCAT	1110
Dd	281	GlyValIleLeuAargIleLeuTyrgluGluasnAspGluSerGluValGluIleIleHis	300
Qy	1111	GTTACATCCCCCTATTGTTGGAAAAACAAGAGGCGCAGATTCAATTCGTTATCTCTAAAAACAGGT	1170
Dd	301	ValThrSerProMetLeuGluThrArgAlaaspSerPheArgTy-ProLysThrGly	320
Qy	1171	ACAGCAAAATCTTAAAGTCACCTTTTAAGATGTCAGAAATAATCATGATGCTCGTAAGAAGG	1230
Dd	321	ThrAlaAsnProLysValThrPheLysMetSerGluIleValLeuASPAlaalaGlyGly	340
Qy	1231	ATCATAGATGCATAGATAAGAACTAAATCAACCTTTTTCAGATTCATATTTGAAGAGATT	1290
Dd	341	IleIleAspValIleApLysgluLeuValGlnProphegluileLeuPhegluGlyVal	360
Qy	1291	GAATATATTGCCAGAGCTGGATGACTCCTGAGGGAAAAATATGCTTGCTCCATCTACTA	1350
Dd	361	GluTyrlleAlaalargAlaGlyTrpThrProGluGlyLysHiseAlaTrpSerilleLeuLeu	380
Qy	1351	GATCGCTCCCAGACTCGCCTACAGATAGTGTTCATCTCACCTGGAATATTTATCCCCAGTA	1410
Dd	381	AspaArgSerGlnThrHisLeuGlnIleValLeuIleserProGluLeuPheIleProVal	400
Qy	1411	GAAGATGATGTTATGAAAGGACAGAGACTCAATGACTCAGTCGCCTGATTTGTGACGCCCA	1470
Dd	401	GluAspAspAlaMetAspArgGlnArgLeuIleGluSerValproAspSer-ValThr-Pro	420
Qy	1471	CYTAATTCATGAAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTT	1530
Dd	421	LeuIlelleTyrgluGluThrThrAspileTrpIleAsnIleHisAspilePheHisVal	440
Qy	1531	TTTTCCCCAAAGTCACGAGAGAAATTGACTTTATTTTTCCTCTGCTGAATGCAAAACAGGT	1590
Dd	441	PheProGlnThrHisGluaspgluileGluPheIlePheAlaserGluCyserThrgly	460
Qy	1591	TTCCTGTCATTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGT	1650
Dd	461	PheArgHisLeutyryLyseIleThrSerilleLeuLysGluSerLysTyryLysArgSerSer	480
Qy	1651	GGTGGCTGCTCTGCCAAGTGATTTCACGTGTCCTATCAAAGAGAGAGATACGAATTACC	1710
Dd	481	GlyGlyLeuProAlaproSerAspPheLysCysProilleLysgluGluIleThrIleThr	500
Qy	1711	AGTGTGTAATGGGAAGTCTTGTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGA	1770
Dd	501	SerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleTrpValAspGluAlaArg	520
Qy	1771	AGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTACAGCATCACCTGTAGCTAGTC	1830
Dd	521	LysLeuValTy-PheGluGlyThrLysAspSerProLeuGluHisIleLeuTyryValThr	540
Qy	1831	AGTTACGTAATTCCTGGAGAGGTGCACAGGCTCAGTCGCGGTACTTCATTTCTTTG	1890
Dd	541	SerTyralaAsnProGlyGluValValargLeuThrAspArgGlyTyry-SerHisSerCys	560
Qy	1891	TGCATCAGTCAGCACTGTGACTCTCTTTTATAGTAAGTATAGTAACACAGAAGAAATCCACAC	1950
Dd	561	CybeLysSerArgHisCysaspPhePheIleserLysTyry-SerAsnGlnLysAsnProHis	580
Qy	1951	TGTGTCTCCCTTTTACAGCTATCAAGTCCTGAAGATGACCCCACTTGC AAAACAAGAA	2010

Db	581	CysValSerLeuTyrLysLeuSerSerProGluAspProValHisLysThrLysGlu	600
QY	2011	TTTTGGGGCCACCATTITGGATTTCAGCAGGTCCTCTCTCTGACTATACCTACCTCTCCAGAAATT	2070
Db	601	PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIle	620
QY	2071	TTCTCTTTTAAAGTACTACTGGATTTACATTGTATGGGATGCTCTACAAGCCTCATGAT	2130
Db	621	PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisasp	640
QY	2131	CTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTCCATATATGTTGGTGCCTCAGGTGAG	2190
Db	641	LeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGln	660
QY	2191	TTGGTGAATAATACGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTACCTAGCCTCTCTA	2250
Db	661	LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeu	680
QY	2251	GGTTATGCGTTGTAGTGTATAGACAACAGGGATCTCTCACCGAGGGCTTAAATTTGAA	2310
Db	681	GlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu	700
QY	2311	GGCGCCTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAGTGGAGGACTCCAA	2370
Db	701	GlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGln	720
QY	2371	TATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGTGGGCATCCAGCGTGTGCC	2430
Db	721	TyrLeuAlaSerGlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPser	740
QY	2431	TATGAGGAGTACTCTCCCTGATGGCATTAATCAGAGGTCAGATATCTTCAGGGTGTCT	2490
Db	741	TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	760
QY	2491	ATTGCTGGGGCCCCAGTCACCTCTCTGGATCTCTATGATACAGGATACACGGAACGTTAT	2550
Db	761	IleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyr	780
QY	2551	ATGGGTACCCCTGCAGCAATGAACAGGGCTATTACTTAGGATCTGTGCCCATCGAACA	2610
Db	781	MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla	800
QY	2611	GAAGAATCCCTCTGTAACCAAAATCGTTTACTGCTCTTACATGGTTCCTGGATGAGAAT	2670
Db	801	GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn	820
QY	2671	GTCCATTTTGCACATACCAGTATATTACTGAGTTTTTTAGTGGGGCTGGAAAGCCATAT	2730
Db	821	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr	840
QY	2731	GATTACAGATCTATCTCTCAGAGAGACACAGCATTAAGAGTTCCTGAATCCGGGAACAT	2790
Db	841	AspLeuGlnIleTyrProGlnLysArgHisSerIleArgValProGluSerGlyGluHis	860
QY	2791	TATGAATCGCATCTTTTCGACTACCTTCAGAAAAACCTTGGATCAGGTATTCCTGCTCTA	2850
Db	861	TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeu	880
QY	2851	AAAGTGATA	2859
Db	881	LysValIle	883

RESULT 10
US-09-976-674-7
; Sequence 7, Application US/09376674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Oi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976,674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-674-7

Alignment Scores:
Pred. No.: 0 Length: 690
Score: 3607.50 Matches: 690
Percent Similarity: 87.1% Conservative: 0
Best Local Similarity: 87.1% Mismatches: 0
Query Match: 65.0% Indels: 102
DB: 3 Gaps: 1

US-10-825-632-2 (1-3120) x US-09-976-674-7 (1-690)

QY	214	ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTGAAACTCCGGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGATATTGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT	333
DB	21	GlulGluIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACAGAAAATATCATGGCTACATGATG	393
DB	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGACCAACATGATGTTTCATGTTGTCAGAGGAGTATGATCCAGATGGACCTCATTC	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAACTTATTACCTTGCATGTCGTGAGACAGAGAAATACACTGTTTATTCT	513
DB	81	AspArgLysTyrTyrLeuAlaMetSerGlyLysAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAAATTCGCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCCTTTTG	573
DB	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGGCAACACTCGACTATGGAATGTTCTTCGAGAGAACTATTAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
QY	634	GAAAGAAAACGATTTGGAACAGATCGGAATTCCTTTACGATTATCACCAAGAGTGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGAGGGCCACAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCGCAACATACGGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATATTATGCGCGCTGATCCAGACTGGATGCTTTTATACATACACAGATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGATATCTAACATCGTACAGAGAAAGGAGCTCACTTATGTGCAACATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLysLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGGAAGAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAA	993

DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
QY	994	TTTCATAGATATTCTGGCTATTGGTGTCTCCAAAGCTGGAACAACTCCAGTGGTGGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAAATCTTAGAATCTTATATGAAGAAATGATGATCTGAGGTGGAATATTATCATCTT	1113
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
QY	1114	ACATCCCTCTATGTTGGAACAACAGGAGGAGATTCATCCGTTATCTCTAAACAGGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATTCATGCTGAGGAGGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGGAACAAATTCACACCTTTTGGAGATTCATTTGAAGGAGTTGAA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATGCCAGAGCTGGATGGACTCTCTGAGGGAATATGCTTGGTCCATCTCTACTAGAT	1353
DB	361	TyrIleAlaAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGTCCAGACTCCCTACAGATAGTGTTCATCTCACCTGAAATTTATTTATCCAGTAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGGAAGGAGAGACTCATTTGAGTTCAGTGCCTGATTCGTGAGCCACTA	1473
DB	401	AspAspValMetGluArgLysArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTCATGTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCAGAAAGAGAAATTTGAGTTATTTTCCCTCTGAAATGCAAAACAGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCAATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAACGATCCAGTGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGGCTGCCTCCCAAGTGATTTCAAGTGCCTATCAAGAGGAGATAGCAATTTACCAGT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGATGGAAATTTCTTGGCGGATCGATCTAATATCCAAAGTTGATGAAGTCAGAAAG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	1834	TACGTAATCTCGAGAGGTGCAAGGCTGACTGACCTGCTGCTACTCACATTTCTTGTGC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGACTGTGACTCTTTTAAAGTATAGTAAGTATAGTACCAAGAGAAATCCACTGT	1953
DB	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCCTTTTAAAGCTATCAAGTCTGAGATGACCAACTCCCAAAACAAAGAAATTT	2013
DB	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCCACTTTTGGATTCAGCAGGTCTCTCTCTGACTATATCTCTCCGAAATTTTC	2073
DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620

QY 2074 TCTTTTGAAGTACTCGATTTACATTTGTATGGATGCTCTACAAGCCTCATGATCTA 2133
Db |||||||
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrglyMetLeuTyrllysProHisAspLeu 640
QY 2134 CAGCCTGGAAGAAATATCCCTACTGCTGTGCTTATATATATGCTGCTCAGGTGCAGTTG 2193
Db |||||||
Db 641 GlnProGlyLyLeLysTyzProThrValLeuPheIleTyrglyGly----- 655
QY 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACCCCTAGCCTCTCTAGGT 2253
Db ----- 655
QY 2254 TATGTGGTTGTAGTAGACACACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGGC 2313
Db ----- 655
QY 2314 GCCTTTAAATATAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAGGACTCCCAATAT 2373
Db ----- 655
QY 2374 CTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGGCTGCTCTAT 2433
Db ----- 655
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTCAGATATCTTCAGGGTGTCTATT 2493
Db |||||-----Arg-LeuLeuLe 659
QY 2494 GCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGATACAGGACGTTATATG 2553
Db |||||||
Db 659 uLeuGlyProGlnSerLeuCysGlySerSerMetIleGlnAspThrArgAsnValIletr 679
QY 2554 GGTCACTCCCTGACCAATGAACAGGGCTATTACT 2587
Db |||||||
Db 679 pValThrLeuThrArgMetAsnArgAlaIleThr 690

RESULT 11
US-10-982-512-7
; Sequence 7, Application US/10982512
; Publication NO. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976, 674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240, 117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-7

Alignment Scores:
Pred. No.: 0 Length: 690
Score: 3607.50 Matches: 690
Percent Similarity: 87.1% Conservative: 0
Best Local Similarity: 87.1% Mismatches: 0
Query Match: 65.0% Indels: 102
DB: 5 Gaps: 1

US-10-825-632-2 (1-3120) x US-10-982-512-7 (1-690)
QY 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGGTGTTGAGATATTTGAAACTGGCGACTGT 273

Db |||||||
1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY |||||||
274 GAGGAGAAATATCGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Db |||||||
21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyzValGluArgTyz 40
QY |||||||
334 TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACAGAAAATATCATGGCTACATGATG 393
Db |||||||
41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyzHisGlyTyzMetMet 60
QY |||||||
394 GCTAAGCCACCATGATTTTCATGTTGTGAAGAGAATGATCCAGATGGACCTCATTC 453
Db |||||||
61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY |||||||
454 GACAGAAATCATTTACCTTCCCATGCTGTGTGTGAGAACAGAGAAAATACACTGTTTATCT 513
Db |||||||
81 AspArgIleTyzTyzLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyzSer 100
QY |||||||
514 GAAATTCCTCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAAAGCCTTTTGG 573
Db |||||||
101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY |||||||
574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAAACTATTAGA 633
Db |||||||
121 AspLeuPheGlnAlaThrLeuAspTyrglyMetTyzSerArgGluGluLeuLeuArg 140
QY |||||||
634 GAAAGAAAACGCAATTCGAAACAGTCGGAATTCCTTACGATTATCACCAAGGAGTGG 693
Db |||||||
141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyzAspTyzHisGlnGlySerGly 160
QY |||||||
694 ACATTTCTGTTCACGCCGTAGTGGAAATTTATACGTAAAGATGAGGAGGCCACAAGGA 753
Db |||||||
161 ThrPheLeuPheGlnAlaGlySerGlyIleTyzHisValLysAspGlyGlyProGlnGly 180
QY |||||||
754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCAACATACGAGTATG 813
Db |||||||
181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY |||||||
814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAAGCATATT 873
Db |||||||
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY |||||||
874 TGGATATCAACATCGTACACAGAGAGAAAGAGACCTCCTTATGTCACAAATGAGCTA 933
Db |||||||
221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyzValHisAsnGluLeu 240
QY |||||||
934 GCCAAATCGAAGAAGATCCAGATCAGCTGGAGTCGCTAGCTTTGTTCTCCAAAGAGAA 993
Db |||||||
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY |||||||
994 TTTGTATAGATATTTCTGGCTATTGTTGGTGTCTCCAAAGCTGAAACAACTCCAGTGGTGT 1053
Db |||||||
261 PheAspArgTyzSerGlyTyzTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY |||||||
1054 AAAATTTCTAGAAATCTATATCAAGAAAATGATGAATCTGAGGTGGGAAAATTTATTCATGTT 1113
Db |||||||
281 LysIleLeuArgIleLeuTyrgluGluAsnAspGluSerGluValGluIleHisVal 300
QY |||||||
1114 ACATCCCTTATGTTGGAAAACAGGGGGCAGATTCATTCCTGTTATCTCTTAAACAGGTACA 1173
Db |||||||
301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyzProLysThrGlyThr 320
QY |||||||
1174 GCAAACTCTAAAGTCACTTTTAAAGTGTCCAGAAAATAATGATTGATGCTGAAGGAGGATC 1233
Db |||||||
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY |||||||
1234 ATAGATGTCATAGATAAGGAACATAATTCAACCTTTTGTAGATTTCTATTTTGAAGGAGTTGAA 1293
Db |||||||
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY |||||||
1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAATATGTTGGTCCATCCTACTAGAT 1353
Db |||||||

361 TyrllealalargalaglyTrpThrProgluGlyLysTyAlaTrpSerlleLeuLeuAsp 380
1354 CGTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATATTATATCCAGTAGAA 1413
381 ArgSerGlnThrArgLeuGlnlleValLeuIleSerProGluLeuPheIleProValGlu 400
1414 GATGATCTTATGAAAGGACGAGACTCATTTGAGTCAGTGCCTGTATCTGTGACCCACTA 1473
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
1474 ATTATCTATCAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT 1533
421 llelleTyrgluGluThrThrAspIleTrrpIleAsnIleHisAspIlePheHisValPhe 440
1534 CCCCAAGTCCAGGAGGAAATTTGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTC 1593
441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGT 1653
461 ArgHisLeuTyrllylleThrSerIleLeuLysGluSerLysTyrllylArgSerSerGly 480
1654 GGGTGGCTGCTCCAGTGATTTCAAGTGTCTCTATCAAGAGAGATAGCAATTACCAGT 1713
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
1714 GGTGAATGGGAAGTCTTGGCCGCATGGATCTTAATTCAGTTCAGTTCAGTTCAGTTCAGT 1773
501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
1774 CTGGTATATTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTACGTACGT 1833
521 LeuValTyrlPheGluGlyThrLysAspSerProLeuGluHisLeuTyrlValSer 540
1834 TACGTAAATCTCGAGAGGTCACAAGGCTGACTGACCGTGGCTGCTGCTGCTGCTGCTGCTG 1893
541 TyrlValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrlSerHisSerCysCys 560
1894 ATCAGTTCAGACTGTGACTCTTCTTATAGTAAGTATAGTAACAGAGAAATCCACACTGT 1953
561 lleSerGlnHisCysAspPhePheIleSerLysTyrlSerAsnGlnLysAsnProHisCys 580
1954 GTGTCCCTTTACAAGCTATCAAGTCTGAAAGATGACCAACTTGCAAAACAAGGAATTT 2013
581 ValSerLeuTyrlLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
2014 TGGGCCACCATTTTGGATTACAGAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrlThrProGluIlePhe 620
2074 TCTTTTGAAGTACTGATTTTACATTTGTATGGGATGCTCTACAGGCTCATGATCTA 2133
621 SerPheGluSerThrThrGlyPheThrLeuTyrlGlyMetLeuTyrlLysProHisAspLeu 640
2134 CAGCCTGGAAAGAAATATCCTACTGCTGCTTCATATATGTTGCTCAGTGCAGTGTG 2193
641 GlnProGlyLysTyrlProThrValLeuPheIleTyrlGly----- 655
2194 GTGAATAATCGGTTTAAAGGAGTCAGTAATTTCCGCTTGAATACCCCTAGCTCTAGGT 2253
655 ----- 655
2254 TATGTGGTTGTAGTAGATAGACAACAGGGGATCTGTCCAGCGGGCTTAAATTTGAAGC 2313
655 ----- 655
2314 GCCTTTAAATATAAAATGGGTCAATAGAAAATTGACGATCAGGTGGAGGACTCCAAATAT 2373
655 ----- 655
2374 CTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGATCCACGCGTGGTCTAT 2433
655 ----- 655

2434 GGAGGATACCTCTCCTGATGGCATTAAATGCAGAGGTCAGATATCTTCAGGCTTGCTATT 2493
656 -----Arg-LeuLeuLe 659
2494 GCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAAGCTTATATG 2553
659 uleuGlyProGlnSerLeuGlySerSerMetIleGlnAspThrArgAsnValIleTr 679
2554 GGTACCTCTGACCAAGATGAACAGGCTATTACT 2587
679 pValThrLeuThrArgMetAsnArgAlaIleThr 690
RESULT 12
US-09-976-674-11
; Sequence 11, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-11
Alignment Scores:
Pred. No.: 1,110-311 Length: 661
Score: 3513.50 Matches: 660
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 63.3% Indels: 2
DB: 3 Gaps: 1
US-10-825-632-2 (1-3120) x US-09-976-674-11 (1-661)
QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCAACAGGATCGCTAAATTTGAGAGCTTTTATTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrlValGluArgIyr 40
QY 334 TCCTGGAGTCAGCTTAAAGGCTGCTTCCGATACCGAAATATCATGCGTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrlHisGlyTyrlMetMet 60
QY 394 GCTAAGGACCAACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTCCCATGCTGTGTGAGAGGAAATGATCCAGATGGACCTCATTC 513
DB 81 AspArgIleTyrlTyrlLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrlSer 100
QY 514 GAAATCCCAAAATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGT 573
DB 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCACTGGACTGGAATGATTTCTCCGAGAGAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrlGlyMetTyrlSerArgGluGluLeuLeuArg 140

Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Qy	274	GAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTGTGAGCGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Qy	334	TCTCGGAGTCAGCTTAAAGCTGCTTGCCTGATACAGAAAATATCATGGCTACATCATG	393
Db	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Qy	394	GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Qy	454	GACGAATCTTATACCTTGCCTGCTGCTGTGTGAGACAGAGAAAATACACTGTTTATTCT	513
Db	81	AspArgIleTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Qy	514	GAATTTCCAAAACCTATCATAGACGAGCAGCTCTTAATGCTCTCTTGGAGCCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTyrLysProLeuLeu	120
Qy	574	GATCTTTTCAGGCAACACTGGACTATGAAATGTTTCTTTCAGTATCATCCAGGAAGTGA	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Qy	634	GAAGAAAACGCATTGGAACAGTCGGAAATTTTCAGTATCATCCAGGAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Qy	694	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTTACGCTAAAAGATGGAGGGCCACAAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTACGCAACACCTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCCAAATATTATGCCCGCTGATCCAGACTGGATGCTTTTATATACATGCAACGATATT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Qy	874	TGGATATCTACATCGTAACAGAGAGAAGAGAGACTCACTTATGTGCACACATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCAACATCGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAAGAAGA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Qy	994	TTTGATAGATATCTGGCTATTGGTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Qy	1054	AAAAATCTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
Qy	1114	ACATCCCTTATGTTGGAACAGAGGGCGAGATTCACTCCGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAATCTTAAAGTCACTTTTAAAGATGTGAGAAAATATGATTTGATGTGAAGAGGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATGTCATAGTAAGGAACCTAATTCACCTTTTGAGATTCCTATTGGAAGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCCAGAGCTGGATGACTCCTCGAGGAAAATATGCTGTGCTCATATATGTTGCTCTC	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaIleTrpSerIleLeuLeuAsp	380
Qy	1354	CGCTCCAGACTCGCTCACAGATAGTGTGATCTCACCTGAATATTATTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	GATCATCTTATGAAAGGAGAGACTCATTGAGTCAGTCCTGATTCTGTGTACGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATGAAAGAAACACAGACATCTGATAAATATCCATGACATCTTTTCATGTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCCAAGTCAAGAAGGAAATAGATTATTTTTCCTCTCTCAATGCAAAACAGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCATTTATACAAAATTTACATCTATTATTAAGGAAGCAATATTAACGATCCAGTGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GGCTGCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAGAGAGATAGCAATATTACAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Qy	1714	GCTGAATCGGAAGTCTTGTGGCGGCATGATCTTAATATCAAGTTGATGAAGTCAGAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTCTTTAGAGCATCACCTGTACGTAGTCACT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
Qy	1834	TACGTAATCTCGAGAGGTGCACAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	1894	ATCAGTCAGACTGTGACTCTCTTATTAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTGTCCTTTTCAAGCTATCAAGTCTGCTGAGAGACCCCACTTGCAAAACAAAGGAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACCATTTTGGATTCAGCAGGCTCTCTTCTGCTACTATCTACTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Qy	2074	TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTTCAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTCGGAAAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCTCCTCCTCCTC	2190
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyLeuLeuArgCysSer	660
Qy	2191	TTGG 2194	
Db	660	ITP 661	
RESULT 14			
US-09-976-674-19			
; Sequence 19, Application US/09976674			
; Patent No. US20020115843A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			

; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Alignment Scores: 8.17e-311 Length: 658
Pred. No.: 3504.00 Matches: 655
Score: 3504.00
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.1% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-19 (1-658)

Qy 214 ATGCGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAGAATATTGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProlysLeuGluProPheTyrValGluArgTyr 40
Qy 334 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCCAGAAATATCATGCTCATCATG 393
Db 41 SerTrpSerGlnLeuIleValLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGCCACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGAGCCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy 454 GACAGAATCTATTACCTTCCCATGCTGTGAGACAGACAGTAAATATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy 514 GAAATTTCCCAAACTATCAATAGACAGCAGCTTAATGCTCTCTCGAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAAGGAAGAACTATTAA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Qy 634 GAAAGAAAACCGCATTCGAAACAGTCGGAATTTGCTTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Qy 694 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATACGTAAAGATGAGGGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Qy 754 TTTACGCAACACTTTAAGGCCAATCTAGTGGAACTAGTTGTCTCCACATACCGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Qy 814 GATCCAAATAATGCCCCCTGATCCAGACTGGATGCTTTTATACATAGCAAGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Qy 874 TGGATATCTAACATCTTAACCCAGAGAAAGAGAGACTCACTTTATGTGCAATAGACTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLysThrTyrValIleAsnGluLeu 240
Qy 934 GCCAATGGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaAlaThrPheValLeuGlnGlu 260

Qy 994 TTTGATAGATATTTCTGGCTATTTGGTGTCTCAAAAGCTGAAACAACTCCAGTGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Qy 1054 AAAATTTCTAGAAATTTATATGAGAAAATGATGATCTGAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Qy 1114 ACATCCCTCTATGTTGGAACAGAGGCGCAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Qy 1174 GCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGCTGAAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Qy 1234 ATAGATGTCATAGATAAGGAACCTAAATTTCAACCTTTTGTAGATTTCTATTTTGAAGGAGTTCAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy 1294 TATATGCCAGAGCTGGATGGAATCTCTGAGGGAATAATGCTGTGCTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Qy 1354 CGTCCGAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Qy 1414 GATGATGTTATGGAAGGCGAGAGCTCATGTTGAGTCAGTCCTGATGCTGTCAGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Qy 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe 440
Qy 1534 CCCCAGAGTCACGAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Qy 1594 CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAATATAAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Qy 1654 GGGCTGCTCTCTCAAGTGTATTTCAGTGTCTCTATCAAGAGGAGATAGCAATACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProLysGluLysGluIleAlaIleThrSer 500
Qy 1714 GGTGAATGGGAAGTTCTTGGCCGCATGATCTTAATTCAGATGATGAGTCAAGAGG 1773
Db 501 GlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Qy 1774 CTGTATATTTTGAAGSCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCACT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
Qy 1834 TACCTAAATCCTGGAGAGGTGACAAGCTGACTGACCGTGGCTACTCACTTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Qy 1894 ATCAGTCAGCACTGTGACTTTCTTTTAAAGTATAGTATATAGTAACAGAGAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy 1954 GTGTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Qy 2014 TGGGCCACCATTTTGGATTCAGCAGGTCTCTCTCTGACTATACCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Qy 2074 TCTTTTGAAGTACTACTGCTGATGTTTACATTTGATGGATGCTCTTACAAGCTCATGATCTA 2133

Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATCTCTACTGTGCTGTTTCATATATGGTGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
RESULT 15
US-10-982-512-19
; Sequence 19, Application US/10982512
; Publication No: US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-19

Alignment Scores:
Pred. No.: 8,17e-311 Length: 658
Score: 3504.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.1% Indels: 0
DB: 5 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-982-512-19 (1-658)

QY 214 ATGGCAGCGCATGGAACAGACAGCTGGTGTGAGATATTGAACTGGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAAATATTGAATCAGAGATCGGCTTAATTTGAGCGCTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACCAAGAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAAATCTATTACCTTCCCATGCTGCTGAGAACAGAGAAATACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTGGAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGGCACACTGGACTATGGAATGATTTCTGAGAGAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCATTGGAACAGTCCGAATTCCTTACGATTATCACCAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAsnSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAGTAAAGATCGAGGCGCAACAGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATTATGCCCGCTGATCCAGACTGGAATGCTTTTATACATGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCGTAACCCAGAGAAAGAGAGACTCACATTATGTCACAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATGGAAGAGATGCCAGATCAGCTGGAGTCCCTACCTTTGCTTCCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaAspSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGTGCTCCAAAGCTGAAACAACTCCCACTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTAGAATTTCTATATGAGAAATGATCAATCTGAGCTGGAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAAACAGAGGCGCAGATTCATTCCTGTTATCCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACATTTTAAGATGTACAGAAATATGATGCTGCTGAGGAAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTTCAGATCTCTATTGGAAGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGAGTGGACTCCTGAGGGAATAATATCTGGTCCATCTCTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCTGAGTGCCTGATTTCTGTGACGCCACTA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGSAAGGCAGAGACTCATTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAAGAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGGTCAAGAGAGAAATTCAGTTTATTTTTCCTGCTGCTGATTCGAAATCCAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAATTCATCTATTTTAAAGGAACAAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTCAATTCAGTGTCTCTATCAAGAGGAGATAGCAATTCACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleHisSer 500
QY 1714 GGTGAATGGGAAGTCTTCTGGCCGATGATCTAATATCCAAAGTTGATGAAGTCAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520

QY 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT 1833
 Db |||||||
 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
 QY 1834 TACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGTACTCACATTCTTGCTGC 1893
 Db |||||||
 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAAGTATAGTAACCCAGAGATCCACACTGT 1953
 Db |||||||
 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 QY 1954 GTGTCCCTTTACAAGCTATCAAGTCTGAAGATGACCCCACTTGCAAAACAAAGGAATTT 2013
 Db |||||||
 581 ValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPhe 600
 QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
 Db |||||||
 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspLysThrProGluLlePhe 620
 QY 2074 TCTTTTGAAGTACTACTGGATTACATTGTATGGGATGCTCTACAGCCCTCATGATCTA 2133
 Db |||||||
 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
 QY 2134 CAGCCTGGAAGAAATATCTCTACTGTGTGCTTCATATATATGGTGGT 2178
 Db |||||||
 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

Search completed: April 14, 2006, 12:42:13
 Job time : 527 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 14, 2006, 12:28:19 ; Search time 9.7 Seconds
(without alignments)
2733.328 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5552
Sequence: 1 aadgtctaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 435010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abss/ABSSWEB spool/US10825632/runat 14042006 092402 3741/app_query.fasta_1
-DB=Published Applications AA New -OFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10825632 @CGN 1.1 32 @runat 14042006 092402 3741 -NCPU=6 -ICPU=3
-NO_WMAP -NGS_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:

1: /SID55/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /SID55/ptodata/2/pubpaa/US06_NEW_PUB pep.*
3: /SID55/ptodata/2/pubpaa/US07_NEW_PUB pep.*
4: /SID55/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
5: /SID55/ptodata/2/pubpaa/US09_NEW_PUB pep.*
6: /SID55/ptodata/2/pubpaa/US10_NEW_PUB pep.*
7: /SID55/ptodata/2/pubpaa/US11_NEW_PUB pep.*
8: /SID55/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	84.7	882	7	US-11-151-601-20
2	651	11.7	745	7	US-11-079-463-6408
3	577	10.4	109	7	US-11-176-951-10
4	551	9.9	627	7	US-11-079-463-7758
5	530	9.5	762	7	US-11-116-939-13
6	529	9.5	738	7	US-11-208-288-4
7	529	9.5	766	6	US-10-501-035-234
8	529	9.5	766	7	US-11-208-288-2
9	525	9.5	766	6	US-10-522-789-2

10	517.5	9.3	760	7	US-11-208-288-6	Sequence 6, Appli
11	454	8.2	760	7	US-11-186-284-55	Sequence 55, Appl
12	394	7.1	99	7	US-11-176-951-11	Sequence 11, Appl
13	255	4.6	624	7	US-11-079-463-7504	Sequence 7504, Ap
14	249.5	4.5	657	7	US-11-179-977-1	Sequence 1, Appli
15	200.5	3.6	737	7	US-11-079-463-9281	Sequence 9281, Ap
16	145	2.6	102	7	US-11-176-951-7	Sequence 7, Appli
17	139	2.5	115	7	US-11-176-951-9	Sequence 9, Appli
18	139	2.5	115	7	US-11-176-951-12	Sequence 12, Appli
19	133.5	2.4	102	7	US-11-176-951-8	Sequence 8, Appli
20	128.5	2.3	1307	6	US-10-995-561-711	Sequence 711, App
21	125	2.3	24	7	US-11-176-951-16	Sequence 16, Appl
22	124.5	2.2	2483	7	US-11-186-999-2	Sequence 2, Appli
23	119.5	2.2	1019	6	US-10-995-561-982	Sequence 982, App
24	119	2.1	710	7	US-11-151-601-232	Sequence 23, Appl
25	118.5	2.1	1155	7	US-11-098-686-10550	Sequence 10550, A
26	118.5	2.1	1178	7	US-11-044-899-29	Sequence 29, Appl
27	118.5	2.1	1243	6	US-10-453-372-1136	Sequence 1136, Ap
28	118	2.1	668	6	US-10-454-437-118	Sequence 118, App
29	118	2.1	1243	6	US-10-453-372-1134	Sequence 1134, Ap
30	118	2.1	4913	6	US-10-453-372-1142	Sequence 1142, Ap
31	118	2.1	4961	6	US-10-453-372-1132	Sequence 1132, Ap
32	115	2.1	849	6	US-10-909-769-18	Sequence 18, Appl
33	113.5	2.0	420	7	US-11-079-463-6314	Sequence 6314, Ap
34	113.5	2.0	2458	7	US-11-186-999-11	Sequence 11, Appl
35	113.5	2.0	2515	7	US-11-113-424-53	Sequence 53, Appl
36	112.5	2.0	1285	7	US-11-206-071-2	Sequence 2, Appli
37	111.5	2.0	1206	6	US-10-995-561-709	Sequence 709, App
38	111.5	2.0	2256	7	US-11-144-368-4	Sequence 4, Appli
39	111.5	2.0	2455	7	US-11-186-999-4	Sequence 4, Appli
40	111.5	2.0	2458	7	US-11-186-999-6	Sequence 6, Appli
41	111.5	2.0	2458	7	US-11-186-999-13	Sequence 13, Appl
42	111	2.0	315	7	US-11-018-868-143	Sequence 143, App
43	110	2.0	1067	7	US/11/062	Sequence 3, Appli
44	110	2.0	1092	7	US/11/062	Sequence 6, Appli
45	110	2.0	2455	7	US-11-186-999-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-11-151-601-20
; Sequence 20, Application US/11515601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtiss, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MP100-054P1RCF10M1D1V1M
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21

```
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 7 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-151-601-20 (1-882)
QY 214 ATGCGACGACGATGGAACAGACAGCTGGGTGTCAGATATTTGAACTCGCGACTGT 273
DB 1 MetAlaAlaMetGlnThrGlnLeuGlyValGluLeuPheGlnThrAlaAspCys 20
QY 274 GAGGAGAAATATGAATCAGAGATCGGCCTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTAAAGAGCTGCTGGCGATACAGAAATATCATCGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATCA 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACCTTGGCATGCTGCTGGTGAGACAGAGAAATACACTGTTTATTCCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAAACTATCAATAGACGACGAGTCTTAATGCTCTCTTGAAGGCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACCATTTGGAACAGTCGGNAITTCCTTACGATTATCACCAAGGAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATTCAGTAAAGATGGAGGGCCACAAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTCGAAACTAGTTGTGCCAACATACGGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATTTATCCCGCTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACATCAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTGCCTTGTCTTCCCAAGAAGAA 993
DB
```

```
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGTAGATATTTCTGGCTATTGCTGCTCCAAAGCTGAAACAACTCCCACTGGTGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGlnThrTrpSerGlyGly 280
QY 1054 AAAATCTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGTGGAAAATTTATCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTCATTTGGAACAAAGAGGCGCAGATTCATTCCTGTTATCTCTAAAACAGGTPACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCTTAAAGTCACATTTTAAAGATGTCAGAAAATAATGATTTGATGCTGAAGGAAGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACTAATTCACACTTTTTCAGATTCCTATTTGAGGAGTTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGCACTCTCGAGGAAAATAATGCTGGTCCATCTACTAGAT 1353
DB 361 TyrIleAlaAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CCGTCCCACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCACTAGAA 1413
DB 381 ArgSerGlnThrArgGluGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGACAGACATCTGAGTAAATATCCATGACATCTTTCATGTTTTT 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAAACAAACAGACATCTGATTAATATCCATGACATCTTTCATGTTTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCAAAGTCACGAAGAGGAAATTCAGTTTATTTTCCCTCTGAATCCAAAACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCAATTTATACAAATTTACATCTATTTTAAAGGAAACAAATAATAACATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGAAAGTTCTTGGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAAG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACTGTCAGTACGTAGTCAGT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuLeuTyrValValSer 540
QY 1834 TACGTAATCTCGAGAGAGTGACAAAGGCTGACTACCGTGGCTACTACATCTCTGCTGC 1893
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTCTTTATATAAGTAAAGTATATACCAAGAGATCCCACTGT 1953
DB 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTCTTTACAAGCTATCAAGTCTCGAAGATCACCACTTCGAAAACAAAGGAATTT 2013
DB 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGGCACCATTTTGGATTCAGCAGGTCTCTCTCTGACTATACCTCTCCAGAAATTTTC 2073
DB 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
```

QY 2074 TCTTTTGAAGTACTACTGGATTACATTGTATGGATGCTCTACAAAGCCTCATGATCTA 2133
DB |||||
QY 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrIysProHisAspLeu 640
DB |||||
QY 2134 CAGCCTGGAAGAAATATCTCTACTGCTGTCTCATATATGATGCTCTCAGGTGCGAGTTG 2193
DB |||||
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB |||||
QY 2194 GTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCAGGT 2253
DB |||||
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB |||||
QY 2254 TATGTGGTGTAGTGATAGACAAACAGGGGATCCTGTCCACGAGGCTTAAATTTGAGGC 2313
DB |||||
QY 681 TyrValValValValLeuAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
DB |||||
QY 2314 GCCTTTAAATATAAATGGTCAAAATAGAAATTCAGCATCAGGTGGAAGACTCCCAATAT 2373
DB |||||
QY 701 AlaPheLysTyrLysMetGlyGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 720
DB |||||
QY 2374 CTAGCTCTCCGATATGATTTTCACTTAGATCGTGTGGCATCCACGCTGGTCTCTAT 2433
DB |||||
QY 721 LeuAlaSerArgTyrAspPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
DB |||||
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGTTGCTATT 2493
DB |||||
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspPheArgValAlaIle 760
DB |||||
QY 2494 GCTGGGGCCGAGTCACTCTCTGTGGATCTTCTATGATACAGGATACAGGAACTTATATG 2553
DB |||||
QY 761 AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMet 780
DB |||||
QY 2554 GGTACCCCTGACCAAGTACAGGCTTATCTAGGATCTGTGGCATTCGACAGGAA 2613
DB |||||
QY 781 GlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
DB |||||
QY 2614 AAGTTCCCTCTGACCAAAATCGTTTACTGCTTACTATGCTTTCTCGATGAGATGTC 2673
DB |||||
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
DB |||||
QY 2674 CATTTTGCACATACAGTATATTTACTGATTTTCTAGTGGGCTGGAAGCCATATGAT 2733
DB |||||
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
DB |||||
QY 2734 TTACAGATCTATCTCTCAGGAGACACAGACATAGAGTTCTCGAATCGGAGAACATTAT 2793
DB |||||
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
DB |||||
QY 2794 GAATGCACTTTTGCATCTACCTCAAGAAACCTTGGATCAGGTATTTGCTCTTAAAA 2853
DB |||||
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
DB |||||
QY 2854 GTGATA 2859
DB |||||
QY 881 ValIle 882
DB |||||

RESULT 2

US-11-079-463-6408
; Sequence 6408, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444

; SEQ ID NO 6408
; LENGTH: 745
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-6408

Alignment Scores: 1.14e-48 Length: 745
Pred. No.: 651.00 Matches: 199
Score: 44.3% Cons: 124
Percent Similarity: 27.3% Mismatches: 244
Best Local Similarity: 11.7% Indels: 162
Query Match: 29
Gaps: 29

US-10-825-632-2 (1-3120) x US-11-079-463-6408 (1-745)

QY 730 GTAAAGATGGAGGCGCACAGGATTTACGCAACACCTTTTAAGGCCCAATCTAGTGAA 789
DB |||||
QY 150 LeuSerAspGlyGlyProGln----- 156
DB |||||
QY 790 ACTAGTTGCCCAACATACGATGGATCCAAATATATGCCCGGTGATCCGACATGGATT 849
DB |||||
QY 157 -----GlnAlaProValPheSerPro---AspGlyAsnLeuVal 168
DB |||||
QY 850 GCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACGAGAGAGAAAGGAGA 909
DB |||||
QY 169 AlaPheValArgAspAsnAsnIlePheLeuVal-----Lys 180
DB |||||
QY 910 CTCACTTATGTGCACAAATGAGTACGCAACATCGGAAGAGATGCCAGA----- 957
DB |||||
QY 181 LeuLeuTyrIysAsnSerGlu---SerGlnValThrGluAspGlyLysLeuAsnSerVal 199
DB |||||
QY 958 TCAGCTGGAGTCTGCTTCTTCTCCAAAGAGAA-----TTTGATAGATATTTCTGGC 1011
DB |||||
QY 200 LeuAsnGlyIleProAspTrpValTyrGluGluGluPheGlyPheAsnArgAlaLeuGlu 219
DB |||||
QY 1012 TATGTGTGTGTCCAAAGCTGAAACAACTCCAGTGTGGTAAATTTCTTAGAATTTCTA 1071
DB |||||
QY 220 Phe-----AsnAlaAspAsnThr-----MetLeuAla 228
DB |||||
QY 1072 TATGAGAAATATGATGATCTGAGGTGGAAATATTTATGTTAGATCCCTATGTG--- 1128
DB |||||
QY 229 TyrValArgPheAspGluSerGluValPro-----SerTyrThrPheProLeuPheAla 246
DB |||||
QY 1129 ---GAAACAGGAGGCGCAGAT-----TCATTCCGTTAT 1158
DB |||||
QY 247 GlyGluAlaProArgTyrAspAlaLeuGlnAspTyrProGlyGluTyrThrTyrLysTyr 266
DB |||||
QY 1159 CCTAAACAGGTACAGCAATCTTAAGTC-----ACTTTT----- 1194
DB |||||
QY 267 ProLysAlaGlyTyrProAsnSerLysValSerValHisThrPheAspIleLysSerLys 286
DB |||||
QY 1195 -----AAGATGTCAAAATATGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGAT 1248
DB |||||
QY 287 ValThrArgGlnValLysLeuProIleAspAlaAspGly----- 299
DB |||||
QY 1249 AAGGAACATATTCACCTTTTGTAGATTTCTTTTGAAGGAGTTGAATATATTTGCCAGAGCT 1308
DB |||||
QY 300 -----TyrIleProArgIle 304
DB |||||
QY 1309 GGATGAGCTCTGAGGGAATAATATGTTGTCCTACTAGATCGCTCCAGACTCGC 1368
DB |||||
QY 305 ArgPheThrGlnAspProAsnLysLeuAlaIleMetThrLeuAsnArgHisGlnAsnArg 324
DB |||||
QY 1369 CTACAGATAGTTGATCTCAGCTGAATTTATTTATCCAGTAGAAGATGATCTTATGAA 1428
DB |||||
QY 325 PheAspMetTyrPheAlaAspPro----- 332
DB |||||
QY 1429 AGCAGAGACTCATTTGAGTCAGTCGCTGATTTCTGTGACGCCCACTAATTTATCTATGAAGA 1488
DB |||||
QY 333 ArgSerThrValCysLysLeuAlaLeuArgAspGluSerProTyrTyrIleAsnGlu---- 351
DB |||||
QY 1489 ACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTACAGAA 1548
DB |||||

Db	352	-----AsnValPheAspAenIleGlnPheProGlu-----	362	660	LysGluAsnAlaGluGlyTyrLysAlaAlaSerAlaPheSerArgAlaAspAsnLeu---	678
Qy	1549	GAGGAAATTCAGTTTATTTTGGCTCTGAATCAAAACAGGTTCCGTCATTATACAAA	1608	2623	TCTGAAACCAATCGTTTACTGCTTACATGGTTTCTCGATGAGAAATGTCCATTTTGCA	2682
Db	363	-----TyrPheSerPheValSerAsp-----LysSerGlyTyrProHisLeuTyr---	377	679	-----HisGlyAsnLeuLeuLeuValHisGlyMetAlaAspAsnValHisPheGln	696
Qy	1609	ATTACATCTATTTTAAAGGAAAGCAATATAACGATCCAGTCGGTGGCTCCCTGCTCCA	1668	2683	CATACCATATATTACTAGTTTCTTACTAGGCTGGGAAGCCATATGATTTTACAGATC	2742
Db	378	-----TyrTyrSerMetAsnGlyAsnLeu-----	385	697	AsnCysThrGluTyrAlaGluHisLeuValGlnLeuGlyLysGlnPheAspMetGlnVal	716
Qy	1669	AGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTACCAGTGGTGAATGGGAAGTT	1728	2743	TATCTCTAGGAGAGACAGACATGAAGTTCCTCGATCCGGGAGAACATTATGACTGCAT	2802
Db	386	-----IleLysGln-----ValThrSerGlyAsnTyrGluVal	396	717	TyrThrAsnArGAsnHisSerIleTyrGlyGlyAsnThrArgAsnHisLeuTyrThrLys	736
Qy	1729	CTTGCGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAAGCGCTGATATTTTGA	1788	2803	CTTTTGCACTACCTTCAAGAAAACCTT	2829
Db	397	LysAsnPheIleGlyTyrAsnProAspThrAsnGlu-----PheTyrTyrThr	412	737	LeuThrAsnPhePheArgAsnAsnLeu	745
Qy	1789	GGCACCAGACTCCCTTTAGAGCATCCTGTAGTAGTCAGTTAGCTAAATCCTGGA	1848	RESULT 3		
Db	413	SerAsnGluSerProMetArgGlnAlaValTyrLysIleAsp-----ArgLysGly	430	US-11-176-951-10		
Qy	1849	GAGGTGACAGGCTGACTACCGCTGCTACTACATCTTGCTGTCATCAGTCAGCACTGT	1908	; Sequence 10, Application US/11176951		
Db	431	LysLysMetLysLeuSerAsnGlnProGlyThrAsnSerProIlePheSerSerSerMet	450	; Publication No. US20060024313A1		
Qy	1909	GACTTCTTTATAAGTATAGTAAACAGGAAATCCACACTGTGTCTCCCTTTACAAG	1968	; GENERAL INFORMATION:		
Db	451	LysTyrPheMetAsnLysPheThrSerLeuAspThrProMetLeuIleThrLeu-----	468	; APPLICANT: CHEN, XIN		
Qy	1969	CTATCAAGTCTGGAAGTACCCCACTTGTCAAACAAAGAAATTTGGGCCACCACTTTTG	2028	; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF		
Db	469	-----AsnAspAsnThrGlyLysValLeuLysThrLeuValThrAsnAsp	483	; FILE REFERENCE: 08842.0019		
Qy	2029	GATTCAGCAGGCTCTCTCTGCTGACTATCTCTCTCCA-----GAAATTTCTCTTTTGA	2082	; CURRENT APPLICATION NUMBER: US/11/176,951		
Db	484	LysLeuLysGlnLysLeuAlaGluTyrAlaIleProGlnLysGluPhePheThrPhelys	503	; PRIOR FILING DATE: 2005-07-06		
Qy	2083	AGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTCGA	2142	; PRIOR FILING DATE: 2004-07-06		
Db	504	ThrThrGluGlyValAspLeuAsnGlyTyrPheMetLysProValAsnPheAspProAla	523	; PRIOR APPLICATION NUMBER: 60/585,095		
Qy	2143	AAGAAATATCTACTGTGCTTTCATATATGTTGGTCTCCTCAGTGCGATGTGTAATAT	2202	; PRIOR FILING DATE: 2004-07-06		
Db	524	LysAspGlyTyrProValLeuMetPheGlnTyrSerGlyProGlySerGlnGlnValLeuAsp	543	; PRIOR APPLICATION NUMBER: 60/585,952		
Qy	2203	CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTATGTGTT	2262	; PRIOR FILING DATE: 2004-07-06		
Db	544	LysTyrPheGlyIleSerTyrPheMetAlaSerLeuGlyTyrValVal	560	; NUMBER OF SEQ ID NOS: 22		
Qy	2263	GTAGTGATAGACACAGGGGCTCTGTCACCGAGGCTTAAATTTCAAGCGCTTTAAA	2322	; SOFTWARE: Patentin Ver. 3.3		
Db	561	AlaCysValAspGlyArgGlyThrGlyGlyArgGlySerGluPheGlnLysCysThrTyr	580	; SEQ ID NO 10		
Qy	2323	TATAAATGGGTCAATAGAAATTCAGCATCAGTCAGGAGGACTCCAATATCTAGCTTCT	2382	; TYPE: PRT		
Db	581	LeuAsnLeuGlyValLysGluAlaLysAspGlnValGluAlaLysTyrLeuGlyGly	600	; ORGANISM: Homo sapiens		
Qy	2383	CGATATGATTCATTGATAGTCGTGGGCTCCACCGCTGCTTATGAGGTTGCTGGGCC	2442	US-11-176-951-10		
Db	601	-----LeuProTyrValAspLysGlyArgIleGlyIleThrGlyTyrPheGlyGlyTyr	619	Alignment Scores:		
Qy	2443	CTCTCCCTGATGGCAATTAATCAGAGGTTCAGATATCTTCAGGTTGCTATTGCTGGGCC	2502	Pred. No.: 1,89e-42		
Db	620	MetThrIleMetSerMetSerMetSerGluGlyThrProValPheLysAlaGlyValAlaValAla	639	Score: 577.00		
Qy	2503	CCAGTCATCTCTGATGATCTTCTATGATACAGGATACACGGAAGTTATATGGGTACCTT	2562	Percent Similarity: 100.0%		
Db	640	AlaProThrAspTyrLysTyrAspThrValTyrThrGluArgPheMetArgThrPro	659	Best Local Similarity: 100.0%		
Qy	2563	GACCAGATGAACAGGGCTATTACTTAGGATCTCTGGCCCATCGCAAGCAAGAAAGTTCC	2622	Query Match: 10.4%		
Db	659	-----	678	DB: 7		
Qy	2622	-----	696	US-10-825-632-2 (1-3120) x US-11-176-951-10 (1-109)		
Db	678	-----	696	Qy 2533 GGATACACGGAACGTTATATGGTTCACCTGACCCAGAAATGAACAGGCTATTACTAGGA		
Qy	2622	-----	696	Db 1 GlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGly		
Db	696	-----	696	Qy 2593 TCTGTGGCCATGCAAGCAGAAAGTCCCTCTGAACCAAAATCGTTTACTGCTCTTACAT		
Qy	2622	-----	696	Db 21 SerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHis		
Db	696	-----	696	Qy 2653 GGTTCCTCGATGAGAAATGTCCATTTTGACATACACAGTATATTTACTGAGTTTTTATGTG		
Qy	2622	-----	696	Db 41 GlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal		
Db	696	-----	696	Qy 2713 AGGCTCGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGATT		
Qy	2622	-----	696	Db 61 ArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgVal		
Db	696	-----	696	Qy 2773 CCTGAATCGGAGAACATTATCACTGCTTTTGGCTACTCTTTCAGAAAAACCTTGA		
Qy	2622	-----	696	Db 81 ProGlnSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGly		
Db	696	-----	696	Qy 2833 TCAGTATTGCTGCTCTAAAGTGATA		
Qy	2622	-----	696	Db 2833 TCAGTATTGCTGCTCTAAAGTGATA		


```
Db      101 SerArgIleAlaAlaLeuLysValIle 109
|||||
RESULT 4
US-11-079-463-7758
; Sequence 7758, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7758
; LENGTH: 627
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-7758

Alignment Scores:
Pred. No.:      6,32e-40      Length:      627
Score:          551.00      Matches:      168
Percent Similarity: 44.1%      Conservative: 120
Best Local Similarity: 25.7%      Mismatches: 249
Query Match:      9.9%      Indels:      116
DB:              7      Gaps:      23

US-10-825-632-2 (1-3120) x US-11-079-463-7758 (1-627)
Qy      916 TATGTGCACAATGAGCTAGCCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACC 975
|||||
Db      76 TyrValAspArgAlaValThrAsnGluProGluGlyIleValCysGlyGlnSer--- 94
|||||
Qy      976 TTTGTCTTCCAAAGAAGATTTGATAGATATTTCTGGCTATTCTGGCTATTGGTGGTCTCCAAAGCTGAA 1035
|||||
Db      95 ---ValHisArgAsnGluPheGlyIleLysLysGlyThrPheTrp----- 108
|||||
Qy      1036 ACACTCCCAAGCTGGTGGTAAATCTTAGAATCTTATATGAGAAATGATGATCTGAG 1095
|||||
Db      109 ---SerProSerGlyAsnLeuLeuAlaPheTyrArgMetAspGlnSerMetValAlaGln 127
|||||
Qy      1096 GTGGAAATTTATCATGTTACATCCCTATGTTGGAAACAAGGGCAGATTCAATCCGT 1155
|||||
Db      128 TyrProLeuValAspValThrAlaProIleAlaGlu-----ValAsnAsnIleArg 144
|||||
Qy      1156 TATCTTAAACAGGTACACCAATCTTAAGTCACTTTTAAGATGTGCAGAAATAATGATT 1215
|||||
Db      145 TyrProMetAlaGlyMetThrSerHisGlnValLysVal-----GlyIleTyrAsn 161
|||||
Qy      1216 GATGCTGAAGGAGGATCATA-----GATGTCATAGATAAGGAACATAATT 1260
|||||
Db      162 ProAlaThrGlyLysSerIleTyrLeuAsnAlaGlyAspProThrAspArg- 178
|||||
Qy      1261 CAACCTTTTGAGATTCATTATTTGAAGAGTTGAATATTTGCCAGAGCTGGATGGACTCCT 1320
|||||
Db      179 -----TyrPheThrAsnIleSerTrpAlaPro 187
|||||
Qy      1321 GAGGGAATAATATCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380
|||||
Db      188 AspGluLysSerLeuTyrLeuIleGluLeuAsnArgAspGlnAsnHisAlaLysLeuCys 207
|||||
Qy      1381 TTGATCTCACCTGAATATTATTATCCAGTAGAAGATGATGTTGGAAAGGCAGAGACTC 1440
|||||
Db      208 ArgTyrAsp----- 210
|||||
Qy      1441 ATTGAGTCAGTCGCTGATTTCTGTGAGCCCACTAATTATCTATGAGAAACAACAGACATC 1500
|||||
```

```
Db      211 ValGluThrGlyGluLeuThrAlaThr-----LeuPheGluGluLysSerAspLys 227
Qy      1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAAGAGCAAATTGAG 1560
|||||
Db      228 TyrValGluProGlnAspProIlePheLeuProTrpAspAsnSer-----Lys 244
|||||
Qy      1561 TTTATTTTTTCCTCTGAATGCAAAAACAGGTTTCCCTCATTTATATACAAAATTTACATCTAT 1620
|||||
Db      245 PheIleTyrGlnSerGln---LysAspGlyPheSerHisLeuTyrLeuTyrAspThrAsn 263
|||||
Qy      1621 TTAAGGAAGAAACAAATATAAACAATCCAGTGTGGTGGCTGCTCCCAAGTGATTCAAG 1680
|||||
Db      264 GlyArgGlnIleArg----- 268
|||||
Qy      1681 TGTCTTATCAAGAGGAGATAGCAATTACCACTGGTGAATGG-----GAAGTTCTT 1731
|||||
Db      269 -----GlnLeuThrGluGlyAspTrpLeuValLysGluValLeu 281
|||||
Qy      1732 GGC CGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGC 1791
|||||
Db      282 Gly-----PheAspThrLysLysLysGluIleIleAlaSer 294
|||||
Qy      1792 ACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAATCTCTGGAGAG 1851
|||||
Db      295 ThrGluPheSerProLeuGlnAsnLeuPheArgLeuAsp---ThrLysThrGlyThr 313
|||||
Qy      1852 GTGACAAGGCTGACTGACCGTGCTACTCACATTTCTGTCATCAGTCAGTCAGCTGTGAC 1911
|||||
Db      314 ArgThrProLeuGlySerAlaGluGlyValHisSerGlyGlnLeuSerProSerGlyArg 333
|||||
Qy      1912 TTCTTTATAAGTAAGTATAGTAACCAAGAAGATCCACACTGTGTGCTCCTTTTACAAAGCTA 1971
|||||
Db      334 TyrLeuIleAspGlnTyrAsnSerProThrValProArgSerIleAsnIleAlaVal 353
|||||
Qy      1972 TCAAGTCTGGAAGATGACCCCACTTCGCAAAACAAAGGAATTTTGGCGCCACCATTTGGAT 2031
|||||
Db      354 GlnSerGlyLysSer-----ValAsnLeuLeuThr 363
|||||
Qy      2032 TCAGCAGGTCTCTTCTGCTACTATATCTCTCTCA-----GAATTTTCTCTTTTCAAGT 2085
|||||
Db      364 AlaAlaAspProPheThrGlyTyrLysMetProGlyIleGluThrGlyThrIleLysAla 383
|||||
Qy      2086 ACTACTGATTTACA---TTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTCGA 2142
|||||
Db      384 AlaAspGlyLysThrAspLeuTyrTyrArgLeuIleLysProAlaAspPheAspProAsn 403
|||||
Qy      2143 AAGAAATATCTACTGTGCTGTTTCATATATATGTTGGTCTCTCAGGTGCTGCTGTAATAT 2202
|||||
Db      404 LysLysTyrProAlaIleValTyrValTyrGlyGlyProHisAlaGlnLeuValThrAsn 423
|||||
Qy      2203 CGGTTTAAA---GGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGTATGCTG 2259
|||||
Db      424 GlyTrpGlnAsnGlyAlaArgGlyTyrAspIle---TyrMetAlaAsnLysGlyTyrIle 442
|||||
Qy      2260 GTTGTAGTATAGACAAACAGGGGATCTCTGTCACCGAGGCTTAAATTTTGAAGGCCCTTT 2319
|||||
Db      443 MetPheThrValAspGlyArgGlySerSerAsnArgGlyLeuAspPheGluAsnValThr 462
|||||
Qy      2320 AAATATAAATGGGTCAATAGAAATTCAGATCAGTGGAGGACTCCAATATCTAGCT 2379
|||||
Db      463 PheArgGlnLeuGlyIleGluGluGlyArgAspGlnValLysGlyThrGluPheLeuLys 482
|||||
Qy      2380 TCTCGATATGATTTTCATTGCTTAGATCGTGGGCATCCAGGCTGCTCTATCGAGGA 2439
|||||
Db      483 Ser---LeuProTyrValAspGlyAsnArgIleGlyValHisGlyTrpSerPheGlyGly 501
|||||
Qy      2440 TACCTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGCTGCTATTCTCTGGG 2499
|||||
Db      502 HisMetThrThrAlaLeuLeuLeuArgTyrProGluIlePheLysValGlyValAlaGly 521
|||||
Qy      2500 GCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCCAC 2559
|||||
Db      522 GlyProValIleAspTrpGlyTyrTyrGluValMetTyrGlyGluArgTyrMetAspThr 541
|||||
```



```
Qy 1567 TTTGCTCTGAATGCAAAACAGGTTTCGTCAT-----TTATAC 1605
Db 349 TrpValGlyArgPheArgProSerGluProHisPheThrLeuAspGlyLeuSerPheTyr 368
Qy 1606 AAAATTTACATCTATTAAAGGAAAGCAAAATATAAAGCATTAACAGATCCAGTGTGGCTGCCTGCT 1665
Db 369 LysIleIleSer-----AsnGluGluGlyTyrArgHisIle----- 380
Qy 1666 CCAAGTGAATTTCAAGTGTCTATCAAGAGAGAGATAGCAATACCAAGTGGTGAATGGAA 1725
Db 381 ---CysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGlu 399
Qy 1726 GTCTTGGCCGCATGGATTAATCAAGTTCATGAAGTCAGAGAGCTGGTGTATATTTT 1785
Db 400 ValIleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrIle 414
Qy 1786 GAAGGCACC---AAAGACTCCCTTTAGAGCATCACCTGTAGCTAGCTAGCTAGTAAT 1842
Db 415 SerAsnGluTyrLysGlyMetProGlyLysArgAsnLeuTyrLysIle----- 430
Qy 1843 CTGGAGAGGTGACAGGCTGACTGACCGCTGCTACTCACATTTCTGCTGCATCACT--- 1899
Db 431 -----GlnLeuSerAsp-----TyrThrLysValThrCysLeuSerCys 443
Qy 1900 -----CAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACAGAGAAATCCA 1947
Db 444 GluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyr 463
Qy 1948 CAC-----TGT-----GTGTCCTTTTCAAGCTCATCAAGCTCTCAAGAT 1986
Db 464 TyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHisSerSerValAsn 483
Qy 1987 GACCAACTTGCAAAACAAAGAAATTTGGGCCACCAATTTGGATTGGATTCAGAGCTCTCTT 2046
Db 484 AspLysGlyLeuArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----Leu 500
Qy 2047 CTGCACTATCTCTCCAGAA-----ATTTTCTCTTTTGAAGTACTACTGATTT 2097
Db 501 GlnAsnValGlnMetProSerLysLysLeuAspPheIleLeuAsnGluThrLysPhe 520
Qy 2098 ACATTTGTATGGATGCTCTACAGGCTCATGATCTACAGCTGGAAGAATATCTACT 2157
Db 521 ---TrpTyrGlnMetIleLeuProHis---PheAspLysSerLysLysTyrProLeu 538
Qy 2158 GTCTGTTTCATATATGGTCTCCTCAGTGCAGTTGTGTAATAATCGTTTAAAGAGATC 2217
Db 539 LeuLeuAspValTyrAlaGlyProCysSerGln-----LysAlaAsp 552
Qy 2218 AGTATTTCCGCTTGAAT-----ACCTAGCTCTCTAGGTTATGTGGTTGTAGTG 2268
Db 553 ThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleIleValAla 572
Qy 2269 ---ATAGACAACAGGGGATCTCTGCACCGAGGCTTAAATTTGAAGCGCTTTAAATAT 2325
Db 573 SerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsnArg 592
Qy 2326 AAAATGGGTCAATAGAAATGACATCAGTGGAGGACTCCCAATATCTAGCTTCTCGA 2385
Db 593 ArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe---SerLys 611
Qy 2386 TATGATTTCAATGACTTAGATCGTGTGGCATCCACGGCTGGTCTATGAGGATACCTC 2445
Db 612 MetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTyrSerTyrGlyTyrVal 631
Qy 2446 TCCTCATGGCATTAAATGACAGGTCAGATATCTTCAGGGTGTCTATGCTGGGATACCTC 2505
Db 632 ThrSerMetValLeuGlySerGlyValPheLysCysGlyIleAlaValAlaPro 651
Qy 2506 GTCACCTGTGGATCTTCTATCATACAGGATACGGAACTTATATGGGT-----CAC 2559
Db 652 ValSerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThr 671
```

2560 CCTGACCAGAATGAACAGGCTATTACTTAGATCTGTGGCCATCGCAAGACGAAAAAGTTTC 2619
672 ProGluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGluAsnPhe 691
2620 CCTCTGGAACCAATGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTCATTTT 2679
692 -----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAspAsnValHisPhe 709
2680 GCACATACCATGATATTACTGAGTTTCTTACTGAGGCTGGAAAGCCATATGATTACAG 2739
710 GlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMet 729
2740 ATCTATCCTCAGAGAGACACAGCATTAAGAGTTCTGAATCGGAGACACATTATGAAC 2799
730 TrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThr 749
2800 CATCTTTTGCACTTACCTTCAAGAA 2823
750 HisMetSerHisPheIleLysGln 757

RESULT 6
US-11-208-288-4
; Sequence 4, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-4

Alignment Scores:
Pred. No.: 5,66e-38 Length: 738
Score: 529.00 Matches: 186
Percent Similarity: 42.0% Conservative: 109
Best Local Similarity: 26.5% Mismatches: 280
Query Match: 9.5% Indels: 128
DB: Gaps: 32

US-10-825-632-2 (1-3120) x US-11-208-288-4 (1-738)

Qy 829 CCCGCTGATCCAGACTGGATT-----GCTTTTATACAT 861
Db 121 ProAsnAsnThrGlnTrpValThrTrpSerProValGlyHisLysLeuAlaTyrValTrp 140
Qy 862 AGCAACGATATTTGGATATCTAACATCGTAACAGAGAGAAAGGAGACTCACTATATGTG 921
Db 141 AsnAsnAspIleTyrValLysIleGluProAsnLeuProSerTyrArgIleThrTrpThr 160
Qy 922 CACAATGAGTCAGCCAACTAGGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTT 981
Db 161 Gly-----LysGluAspIleIleTyrAsnGlyIleThrAspTrpVal 174
Qy 982 CTCACAAGAA-----TTTGATGATATTTCTGCTATTGCTGTGTCCTCAAAAGCTGAAACA 1038
Db 175 TyrGluGluGluValPheSerAlaTyrSerAlaLeuAlaLeuTrpSerProAsnGlyThrPhe 194
Qy 1039 ACTCCAGTGGTGAATAATTTCTAGAAATTTCTATGAGAAATGATGAATCTCTGAGGTG 1098
Db 195 -----LeuAlaTyrAlaGlnPheAsnAspThrGluVal 205
Qy 1099 GAAATTATTCAT-----GTTACATCCCTATGTTGGAACAACAGAGGCGCAGATCA 1149

Db 739 AspHisGlyLeuAlaSerSerThrAlaHisGlnHisIleTyrThrHisMetSerHisPhe 758
Qy 2815 CTTCAAGAA 2823
Db 759 IleLysGln 761
RESULT 8
US-11-208-288-2
; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-2

Alignment Scores:
Pred. No.: 5,72e-38 Length: 766
Score: 529.00 Matches: 186
Percent Similarity: 42.0% Conservative: 109
Best Local Similarity: 26.5% Mismatches: 280
Query Match: 9.5% Indels: 128
DB: 7 Gaps: 32

US-10-825-632-2 (1-3120) x US-11-208-288-2 (1-766)

Qy 829 CCCGCTGATCCAGACTCGATT-----GCTTTTATACAT 861
Db 149 ProAsnSerThrGlnTrpValThrTrpSerProValGlyHisLysLeuAlaTyrValTrp 168
Qy 862 AGCAACGATATTGGATATCTAACATCGTAAACAGAGAGAAGAGAGACTCACTTATGTG 921
Db 169 AsnAsnAspIleTyrValLysIleGluProAsnLeuProSerTyrArgIleThrTrpThr 188
Qy 922 CACAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGAGTGCCTACCTTTGTT 981
Db 189 Gly-----LysGluAspIleIleTyrAsnGlyIleThrAspTrpVal 202
Qy 982 CTCCAAGAAGAA-----TTTGATAGATATTCTGGCTATTGGTGTGTCCTCAAAAGCTGAACA 1038
Db 203 TyrGluGluGluValPheSerAlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe 222
Qy 1039 ACTCCAGTGTGTGTAATAATCTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTG 1098
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluVal 233
Qy 1099 GAAATTATTCAT-----GTTACATCCCTATCTTGAACAACAGAGAGGCAGATTCA 1149
Db 234 ProLeuIleGluTyrSerPheTyrSerAspGluSerLeuGlnTyrProLysThrValArg 253
Qy 1150 TTCGGTATCTTAAACAGATACAGCAATCTTAAGTCACTTTAAGATGTCAGAAATA 1209
Db 254 ValProTyrProLysAlaGlyAlaValAsnProThrValLysPheVal-----Val 271
Qy 1210 ATGATTTGATGCTGAAGAGATCATAGATGTCATAGATAAGAACTAATTCACCTTTT 1269
Db 272 AsnThrAspSerLeuSerValThrAsnAlaThrSerIleGlnIleThrAlaProAla 291
Qy 1270 GAGATTCTATTGAAGAGATTGAATATTGTCAGAGCTGGATGGACTCCTGAGGGAATA 1329
Db 292 SerMetLeuIle-----GlyAspHisTyrLeuCysAspValThrTrpAla----- 306

Qy 1330 TATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTAGTGTGATCTCA 1389
Db 307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu--- 316
Qy 1390 CTTGAATTATTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTTAGTCA 1449
Db 317 -----ArgArgIleGlnAsn 321
Qy 1450 GTGCTCGATTCTGTGAGCGCCACTAATTATCTATGAAGAAACACAGACATCTCGATAAAT 1509
Db 322 Tyr-----SerValMetAspIleCysAspTyrAspGluSerSerGlyArgTrp---Asn 338
Qy 1510 ATCCATGACATCTTTTCATGTTTTTCCCAAAAGTCCAGAAAGGAGAAATGAGTTATTTT 1569
Db 339 Cys-----LeuValalaArgGlnHisIleGluMetSerThrThrGlyTrp 353
Qy 1570 GCCTCTGAATGCAAAACAGGTTTCCGTCAT-----TTATACAAA 1608
Db 354 ValGlyArgPheArgProSerGluProHisPheThrLeuAspGlyAsnSerPheTyrLys 373
Qy 1609 ATTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGTGGGCTGCCTGCTCCA 1668
Db 374 IleIleSer-----AsnGluGluGlyTyrArgHisIle----- 384
Qy 1669 AGTGATTTCAAGTGTCTTATCAAGAGAGATAGCAATTACCATGCTGTGAATGGGAAGTT 1728
Db 385 CysTyrPheGlnIleAspLysAspCysThrPheIleThrLysGlyThrTrpGluVal 404
Qy 1729 CTTGGCGCGCATGGATCTAATATCAAGTTGATGAAGTCAGAAAGCTGGGTATTTTGA 1788
Db 405 IleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSer 419
Qy 1789 GGCACC---AAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAATCCT 1845
Db 420 AsnGluTyrLysGlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyr 439
Qy 1846 GGAGAGGTGACAAAGCTGACTGACCGCTGCTACTCACATTTCTTCTGCTCATCAGT---CAG 1902
Db 440 ThrLysValThrCysLeu-----SerCysGluLeuAsnProGlu 452
Qy 1903 CACTGTGATCTCTTAAAGTATAGTATAGTAAACCAAGAAATCCACAC-----TGT 1953
Db 453 ArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCys 472
Qy 1954 -----GTGCTCCTTTACAGCTATCAAGTCTCAAGATGACCCCACTTGCAAA 2001
Db 473 SerGlyProGlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArg 492
Qy 2002 ACAAGAGAAATTTGGGCCACCATTTTGGATTCAGAGGCTCCTCTTCTGCTACTACTCTCT 2061
Db 493 ValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMet 509
Qy 2062 CCAGAA-----ATTTTCTTTTGAAGTACTACTGTGATTTACATTGTATGGGATG 2112
Db 510 ProSerLysLysLeuAspPheIleLeuAsnGluThrLysPhe---TrpTyrGlnMet 528
Qy 2113 CTCCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTGCTGTTTCATATAT 2172
Db 529 IleLeuProProHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyr 547
Qy 2173 GGTGGTCTCAGGTGCTGTGTAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTG 2232
Db 548 AlaGlyProCysSerGln-----LysAlaAspThrValPheArgLeu 561
Qy 2233 AAT-----ACCTAGCTCTCTAGTTATCTGGTGTGTAGTG---ATAGACAAACAGG 2280
Db 562 AsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArg 581
Qy 2281 GGATCTGTCCAGGAGGCTTAATTTGAAGCGCCCTTAAATATAATAAATGGGTCAATA 2340
Db 582 GlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrPhe 601
Qy 2341 GAAATTGACATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATGATGTTTCATTGAC 2400

```
Db 602 GluValGluAspGlnIleGluAlaAaArgGlnPhe---SerLysMetGlyPheValAsp 620
QY 2401 TTAGATCGTGTGGGATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
Db 621 AsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGlyTyrValThrSerMetValLeu 640
QY 2461 ATGCAGAGGTGAGATATCTTCAGGGTTGCTATGCTGGGGCCCGAGTCACTCTGTGGATC 2520
Db 641 GlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGlu 660
QY 2521 TTCTATGATACAGGATACACGGAACGTTATATGGGT-----CACCTGCACGAGATGAA 2574
Db 661 TyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThrProGluAspAsnLeu 680
QY 2575 CAGGGCTATTACTTAGGATCTGTGGCCATCGACAGCAGAAAGTTCCCTCTGAACCAAT 2634
Db 681 AspHisTyrArgAsnSerThrValMetSerArgAlaGluAsnPhe-----LysGlnVal 698
QY 2635 CGTTTACTGCTCTTACATGGTTCCCTGGATGAGATGTCCATTTTGCACATACCGATATA 2694
Db 699 GluTyrLeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGln 718
QY 2695 TTACTGAGTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCCTCAGGAG 2754
Db 719 IleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMetTyrTrpThrAspGlu 738
QY 2755 AGACACACATAAAGTTCTCGAATCGGAGAACATTTAGAACTGTCATCTTTTGGCACTAC 2814
Db 739 AspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThrHisMetSerHisPhe 758
QY 2815 CTTCAAGAA 2823
Db 759 IleLysGln 761
```

RESULT 9

```
US-10-522-789-2
; Sequence 2, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2
```

Alignment Scores:

Pred. No.:	1-28e-37	Length:	766
Score:	525.00	Matches:	186
Percent Similarity:	41.6%	Conservative:	108
Best Local Similarity:	26.3%	Mismatches:	277
Query Match:	9.5%	Indels:	136
DB:	6	Gaps:	33

US-10-825-632-2 (1-3120) x US-10-522-789-2 (1-766)

```
QY 829 CCGCTGATCCAGACTGGATT-----GCTTTTATACAT 861
Db 149 ProAsnAsnThrGlnTrpValThrTrpSerProValGlyHisLysLeuAlaTyrValTrp 168
QY 862 AGCAACGATATTGGATCTCTAACATCGTAAACAGAGAGAAAGGAGACTCATCTTATGCG 921
Db 169 AsnAsnAspIleTyrValLysIleGluProAsnLeuProSerTyrArgIleThrTrpThr 188
```

```
QY 922 CACAATGAGCTAGCCAAACATGGAAGACATGCCAGATCAGCTGGAGTCGCTACCTTGTGT 981
Db 189 Gly-----LysGluAspIleIleTyrAsnGlyIleThrAspTrpVal 202
QY 982 CTCCAAGAAAGAA---TTTGATAGATATTTCTGGCTATTGGTGTGTCACAAAGAGCTGAACA 1038
Db 203 TyrGluGluGluValPheSerAlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe 222
QY 1039 ACTCCCAGTGGTGAATAATCTTAGAATTCATATGAAGAAATGATGAATCTGAGGTG 1098
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluVal 233
QY 1099 GAAATTTATTCAT-----GTTACATCCCCCTATGTTGGAACAAGAGGAGCGAGATTCA 1149
Db 234 ProLeuIleGluTyrSerPheTyrSerAspGluSerLeuGlnTyrProLysThrValArg 253
QY 1150 TTCCTGTTATCTTAAACACAGGTACAGCAATCCTAAAGTCACCTTTTAAAGATCTCAGAATA 1209
Db 254 ValProTyrProLysAlaGlyAlaValAsnProThrValLysPhePheVal-----Val 271
QY 1210 ATGATTGATGCTGGAAGGAAGGATCATAGATGTCATAGATAGGAACATAATTCACACCTTTT 1269
Db 272 AsnThrAspSerLeuSerValThrAsnAlaThrSerIleGlnIleThrAlaProAla 291
QY 1270 GAGATTCTATTGAAGGAGTTGATATATATATGCGCAGCTGAGTGGAGTCTCTCAGGAGAAA 1329
Db 292 SerMetLeuIle---GlyAspHisTyrLeuCysAspValThrTrpAla----- 306
QY 1330 TATGCTTGGTCCATCTACTAGATCGCTCCCAAGTCTGCTACAGATAGTGTGATCTCA 1389
Db 307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu--- 316
QY 1390 CCTGAATTATTATCCCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCATTTAGATCA 1449
Db 317 -----ArgArgIleGlnAsn 321
QY 1450 GTGCTGATTTCTGTGACGCCACTTAATTATCTATGAAAGAAACAACAGACATCTGGATAAAT 1509
Db 322 Tyr-----SerValMetAspIleCysAspTyrAspGluSerSerGlyArgTrp---Asn 338
QY 1510 ATCCATGACATCTTTTCAATGTTTCCCAAGTCTCAGAGAGGAAATTTAGTTATTTT 1569
Db 339 Cys-----LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrp 353
QY 1570 GCCTCTGAATGCAAAACAGTTTCCGTCAT-----TTATACAAA 1608
Db 354 ValGlyArgPheArgProSerGluProHisPheThrLeuAspGlyAsnSerPheTyrLys 373
QY 1609 ATTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCTCCA 1668
Db 374 IleIleSer-----AsnGluGluGlyTyrArgHisIle----- 384
QY 1669 AGTGATTTCAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTT 1728
Db 385 CysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGluVal 404
QY 1729 CTTGGCGGCATCGATCTAATATCCAAAGTTGATGAAGTCAAGAGCTGTATATTTGAA 1788
Db 405 IleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSer 419
QY 1789 GGCACC---AAAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTACGTTACGTAATCCT 1845
Db 420 AsnGluTyrLysGlyMetProGlyGlyArgAsnLeuTyrLysIle----- 434
QY 1846 GGAGAGGTGACAAGGCTGACTGACCTGGCTACTACATTTCTTGTGTCATCAGT----- 1899
Db 435 -----GlnLeuSerAsp-----TyrThrLysValThrCysLeuSerCysGlu 448
QY 1900 -----CAGCACTGTGACTCTCTTTTATAAGTAGTATAGTAACCAAGAAATCCACAC 1950
Db 449 LeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyr 468
QY 1951 -----TGT-----GTGCTCCCTTTTACAAGCTATCAAGTCTCTCAAGATGAC 1989
```

```
Db 469 GlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHisSerSerValAsnAsp 488
;
Qy 1990 CCAACTTGCACAAACAGGAATTTGGCCACCATTTGGATTTCAGCAGGCTCTTCTCT 2049
;
Db 489 LysGlyLeuArgValLeuGluAsp---AenSerAlaLeuAspLysMet-----LeuGln 505
;
Qy 2050 GACTATACTCTCCAGAA-----ATTTCCTCTTTTGAAGTACTACTGATTTACA 2100
;
Db 506 AenValGlnMetProSerLysLysLeuAspPheLeileLeuAenGlnThrLysPhe--- 524
;
Qy 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT 2160
;
Db 525 TrpTyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrProLeuLeu 543
;
Qy 2161 CTGTTCATATATGTTGCTCTCAGGTCGATTTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
;
Db 544 LeuAspValTyrAlaGlyProCysSerGln-----LysAlaAspThr 557
;
Qy 2221 TATTTCCGCTTGAAT-----ACCTAGCCTCTCTAGTTATGTGTGTGTAGTG--- 2268
;
Db 558 ValPheArgLeuAenTyrPalatThrTyrLeuAlaSerThrGluAenIleileValAlaSer 577
;
Qy 2269 ATAGACACAGGGATCTGTCCACGAGGCTTAAATTTGAAGGCGCTTTTAAATATAA 2328
;
Db 578 PheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAenArgArg 597
;
Qy 2329 ATGGGTCAAATAGAAATTCAGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATAT 2388
;
Db 598 LeuGlyThrPheGluValGluAspGlnIleGluAlaArgGlnPhe---SerLysMet 616
;
Qy 2389 GATTCATTGATTTAGATCTGTGGGATCCAGGCTGCTCTATGGAGGATACCTCTCC 2448
;
Db 617 GlyPheValAspAenLysArgIleAlaIleTrpGlyTrpSerTyrGlyTyrValThr 636
;
Qy 2449 CTGATGGCATTAATCAGAGGTGATATCTTCAGGTTGCTATGCTGGGCGCCAGTC 2508
;
Db 637 SerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProVal 656
;
Qy 2509 ACTCTGTGATCTCTATGATACAGGATACACGGAACTGTTATATGGGT-----CACCT 2562
;
Db 657 SerArgTrpGluTyrTyrAspSerValTyrGluArgTyrMetGlyLeuProThrPro 676
;
Qy 2563 GACCAGAAATGACAGGCTATTACTAGGATCTGTGGCCATGTCAGCAGAGAAAGTTCCCC 2622
;
Db 677 GluAspAenLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGluAenPhe--- 695
;
Qy 2623 TCTGAACCAATCGTTTACTCTCTTACATGTTTCTCGATGAGAAATGTCATTTGCA 2682
;
Db 696 ---LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGln 714
;
Qy 2683 CATACCGATATATTACTGAGTTTATTAGTGGGCTGGAAAGCCATATGATTTACAGATC 2742
;
Db 715 GlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMetTrp 734
;
Qy 2743 TATCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTATGAATGAT 2802
;
Db 735 TyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThrHis 754
;
Qy 2803 CTTTGTGCACTACCTTCAAGAA 2823
;
Db 755 MetSerHisPheIleLysGln 761
```

RESULT 10

```
US-11-208-288-6
; Sequence 6, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
```

```
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-208-288-6

Alignment Scores:
Pred. No.: 5,828-37 Length: 760
Score: 517.50 Matches: 203
Percent Similarity: 38.1% Conservative: 107
Best Local Similarity: 25.0% Mismatches: 306
Query Match: 9.3% Indels: 197
DB: 7 Gaps: 37

US-10-825-632-2 (1-3120) x US-11-208-288-6 (1-760)
Qy 604 ATGTATTTCTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATTGGAAACAGTCGGAATT 663
;
Db 67 LeuTyrLysGlnGluAenAsnIleLeuLeuLeuAsnAlaGluHisGlyAsnSerSerIle 86
;
Qy 664 -----GCTTCTTACGATTATCACCAGAGGAGTGGAAACATTTCTG 702
;
Db 87 PheLeuGluAenSerThrPheGluSerPheGlyTyrHisSerValSerProAspArgLeu 106
;
Qy 703 TTT-----CAAGCCGGTAGT 717
;
Db 107 PheValLeuLeuGluTyrAsnTyrValLysGlnTrpArgHisSerTyrThrAlaSerTyr 126
;
Qy 718 GGAATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGCAACAACCTTTAAGGCC 777
;
Db 127 AsnIleTyrAspValAsnLys-----ArgGln 135
;
Qy 778 ATCTAGTGGAACTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGAT 837
;
Db 136 LeuIleThrGluGluLysIleProAsn-----Asn 145
;
Qy 838 CCAGACTGGATT-----GCTTTATACATACGCAACGAT 870
;
Db 146 ThrGlnTrpIleThrTrpSerProGluGlyHisLysLeuAlaTyrValTrpLysAsnAsp 165
;
Qy 871 ATTTGATATCTAACATCGTAAACAGAGAAAGAGACTCCTATTATGTGCACAATGAG 930
;
Db 166 IleTyrValLysValGluProHisLeuProSerHisArgIleThr-----180
;
Qy 931 CTAGCCCAACATCGAAGAGATGCCAGATCAGCTGAGTGCCTACTTTGTTCTCAAGAA 990
;
Db 181 ---SerThrGlyGluGluAenValIleTyrAsnGlyIleThrAspTrpValTyrGluGlu 199
;
Qy 991 GAA---TTTGTAGATATTTCTGGCTATTGGTGTGTCACAAA-----1029
;
Db 200 GluValPheGlyAlaTyrSerAlaLeuTrpTrpSerProAenAenThrPheLeuAlaTyr 219
;
Qy 1030 GCTGAAACAACTCCAGTGGTGTAAATCTTGAATTTCTATATGAAGAAATGATGAA 1089
;
Db 220 AlaGlnPheAsnAspThrGlyValProLeuIleGluTyrSerPhe---TyrSerAspGlu 238
;
Qy 1090 TCTGAGGTGGAATATTATTCATGTTACATCCCTATGTTGGAACAAGAGGCGGAGATTC 1149
;
Db 239 Ser-----LeuGlnTyrProLysThrValTrp 247
;
Qy 1150 TTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAATA 1209
;
Db 248 IleProTyrProLysAlaGlyAlaValAenProThrValLysPhePheIle-----Val 265
;
Qy 1210 ATGATTGATGCTGAAGGAGGATCATAGTGTCTAGATAGGAACTAATTAACCTTTT 1269
;
Db 266 AsnIleAspSerLeuSerSerSerSerSerAlaAlaProIleGlnIleProAlaProAla 285
```


Alignment Scores:			
Pred. No.:	2,15e-31	Length:	760
Score:	454.00	Matches:	169
Percent Similarity:	40.2%	Conservative:	113
Best Local Similarity:	24.1%	Mismatches:	269
Query Match:	8.2%	Indels:	150
DB:	7	Gaps:	27
US-10-825-632-2 (1-3120) x US-11-186-284-55 (1-760)			
Qy	847	ATTCTCTTTTATACATGACACGATATTTGGATATCTACATCGTAACACAGAGAAGAAGG	906
Db	162	LeuAlaTyRValTyRGlnAAsnIleTyRLeuLysGlnA-gProGlyAspProPhe	181
Qy	907	AGACTCACTTATGTGCACATGAGCTAGCCACATGGAAGAGATGCCAGATCAGCTGGA	966
Db	182	GlnIleThrPhe-----AAsnGlyA-gGluAAsnLysIlePheAAsnGly	195
Qy	967	GTCGCTACCTTTGTTCTCCAAAGAATTT-----GATGATATTTCTGGCTATTGGTGG	1020
Db	196	IleProAspTrpValTyRGluGluMetLeuProThrLysTyRAla---LeuTrpTrp	214
Qy	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATCTTATATGAAGAA	1080
Db	215	SerPro-----AAsnGlyLysPheLeu-----AlaTyRAlaGlu	225
Qy	1081	AATGATGAATCTGAGTGGAATTTATTCATGTTTACATCCCTCTATGTTGGAACA---AGG	1137
Db	226	PheAAsnLysAspIleProValIleAlaTyR-SerTyR-TyR-GlyAspGluGlnTyRPro	245
Qy	1138	AGGCGAGATTCATCCGTTATCTTAAACAGGTACAGCAATCCCTAAAGTCACTTTTAAAG	1197
Db	246	ArgThrIleAAsnIleProTyRProLysAlaGlyAlaLysAsnProValValArgIle---	264
Qy	1198	ATGTCAGAAATATGATTGCTGGAAGGAGATCATAGATCTCATAGATGAAGAACTA	1257
Db	265	-----PheIleAspThrThrTyRProAlaTyR	274
Qy	1258	ATTCAACCTTTTGGAGATT-----CTATTGAGGAGTTGAATATATTATTC	1302
Db	275	ValGlyProGlnGluValProValProAlaMetIleAlaSerSerAspTyR-TyrPheSer	294
Qy	1303	AGAGCTGGATGACCTCTGAGGGAATAATATGCTTGGTCCATCTCATAGATCGCTCCGAG	1362
Db	295	TrpLeuThrTrpValThrAspGluArgValCysLeuGlnTrpLeu-----	309
Qy	1363	ACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTGAAGATGATGTT	1422
Db	310	LysArgValGlnAAsnValSerValLeuSerIleCysAspPheArgGluAspTrpGlnThr	329
Qy	1423	ATGGAAGCGACAGACTCATTCAGTCAGTCAGTCTGATTCCTGTCGACG-----	1467
Db	330	TrpAspCysProLysThrGlnGluHisIleGluGluSerArgThrGlyTrpAlaGlyGly	349
Qy	1468	-----CCACTAATTATCTATGAAGAAACACAGACATCTCGATAAATATC	1512
Db	350	PhePheValSerArgProValPheSerTyRAspAla-----IleSerTyR	364
Qy	1513	CATGACATCTTTTCATGCTTTTCCCAAAGTCACGAGAGAAATGAGTTTATTTTGGC	1572
Db	365	TyrLysIlePhe-----	368
Qy	1573	TCTGAATGCAAAACAGGTTTCGTCATTATACAAAATTTACATCTATTTTAAAGAAAGC	1632
Db	369	---SerAspLysAspGlyTyRHisIleHisTyRile-----	380
Qy	1633	AAATATAACGATCCAGTGGTGGCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAA	1692
Db	381	-----LysAspThrValGlu	385
Qy	1693	GAGGAGATAGCAATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCGCATGATCTAATATC	1752

Db	386	AsnAlaIleGlnIleThrSerGlyLysTrpGluAlaIle-----AAsnIle	400
Qy	1753	CAAGTTGATGAAGTCAGAAAGCTGGTATATATTTTGAAGGCACCAAGAACTCCCTTTAGAG	1812
Db	401	-----PheArgValThrGlnAspSerLeuPheTyR	410
Qy	1813	CATCACCTGTACGTAGTACAGTTACGTAAATCTGGAGAGGTGACAGCTGACGCGT	1872
Db	411	SerSerAsnGluPheGluGluTyRProGlyArgArgAAsnIleTyRArgIleSerIleGly	430
Qy	1873	GGCTACTACATTTCT---TGCTGCATCAGTCAGCAC-----TGTCACCTC	1914
Db	431	SerTyRProProSerLysCysValThrCysHisLeuArgLysGluArgCysGlnTyR	450
Qy	1915	TTTATAAGTAAATATAGTAAACAGAGAATCCACACTGTGTGTCCTTTAC-----	1965
Db	451	TyrThrAlaSerPheSerAspTyRAlaLysTyRTrpAlaLeuValCysTyR-GlyProGly	470
Qy	1966	---AAGCTATCAAGTCTGAGATGACCAACTCTGCAAAACAAAGAAATTTTGGGCCACC	2022
Db	471	IleProIleSerThrLeuHisAspGlyArgThrAspGlnGlu-----IleLys	486
Qy	2023	ATTTTGGATTCAGCAGGTCTCTCTCTGAC-----TATACTCTCTCCAGAA	2067
Db	487	IleLeuGluGluAAsnLysGluLeuGluAAsnIleGlnLeuProLysGlu	506
Qy	2068	ATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCTCAT	2127
Db	507	GluIleLysLysLeuGluValAspGluIleThrLeuTyR-TyR-LysMetIleLeuProPro	526
Qy	2128	GATCTACAGCTGGAAGAAATATCTCTACTGTCTCTCATATATATGGTGTCTCAGGTG	2187
Db	527	GlnPheAspArgSerLysTyRProLeuLeuIleGlnValTyR-GlyGlyProCysSer	546
Qy	2188	CAGTTGGTGAATATCCGTTTAAAGGAGTCAAGTATTTCCGTTGATACCTACGCTCT	2247
Db	547	GlnSerValArgSer-----ValPheAlaValAAsnTrpIleSerTyR	560
Qy	2248	CTA-----GGTTATGTGGTTGTAGTAGACACACAGGGGATCTGTCCCGA	2295
Db	561	LeuAlaSerLysGluGlyMetValIleAlaLeuValAspGlyArgGlyThrAlaPheGln	580
Qy	2296	GGGCTTAAATTTGAAGCGCTTTTAAATATAAATGGGTCAAATAGAAATTCAGATCAG	2355
Db	581	GlyAspLysLeuLeuTyRAlaValTyRArgLysLeuGlyValTyRGluValGluAspGln	600
Qy	2356	GTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTCAGTTAGATCGTGGGC	2415
Db	601	IleThrAlaValArgLysPheIle---GluMetGlyPheIleAspGluLysArgIleAla	619
Qy	2416	ATCCAGCTGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATTCAGAGGTGAGT	2475
Db	620	IleTrpGlyTrpSerTyR-GlyGlyTyRValSerSerLeuAlaLeuAlaSerGlyThrGly	639
Qy	2476	ATCTTCAGGTTGCTATTGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGATACAGGA	2535
Db	640	LeuPheLysCysGlyIleAlaValAlaProValSerSerTrpGluTyRTrpAlaSerVal	659
Qy	2536	TACACGGAACCTTATATGGTCACCT-----GACCAGATGAACAGGCTATTACTTA	2589
Db	660	TyrThrGluArgPheMetGlyLeuProThrLysAspAspAsnLeuGluHisTyR-LysAsn	679
Qy	2590	GGATCTGTGGCCATGCACAGCAAAAAGTTCCTCTGACCAATCTGTTACTGCTCTTA	2649
Db	680	SerThrValMetAlaArgAlaGluTyR-PheArgAsnValAsp-----TyrLeuLeuIle	697
Qy	2650	CATGGTTCTTCGATGAGAAATGTCATTTTGCATACCATCAGTATATATCTAGTTTTTA	2709
Db	698	HisGlyThrAlaAspAsnValHisPheGlnAAsnSerAlaGlnIleAlaLysAlaLeu	717
Qy	2710	GTGAGGCTGGAAGCCATATGATTTACAGATCTTATCTCTCAGGAGACAGCAGATAGA	2769
Db	718	ValAAsnAlaGlnValAspPheGlnAlaMetTrpTyR-SerAspGlnAAsnHisGlyLeu---	736

```
QY 2770 GTTCTGAATCGGGA-----GAACATTATGAACCTGTCATCTTTTGGCACTACCTTCAA 2820
Db 737 -----SerGlyLeuSerThrAsnHisLeuTyrThrHisMetThrHisPheLeuLys 753
QY 2821 GAA 2823
Db 754 Gln 754

RESULT 12
US-11-176-951-11
; Sequence 11, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-11

Alignment Scores:
Pred. No.: 2,03e-26 Length: 99
Score: 394.00 Matches: 72
Percent Similarity: 82.8% Conservative: 10
Best Local Similarity: 72.7% Mismatches: 17
Query Match: 7.1% Indels: 0
Gaps: 0

US-10-825-632-2 (1-3120) x US-11-176-951-11 (1-99)
QY 2533 GGATACCGGAAGCTTATATGGTCACTGACAGATGAACAGGCTATTACTTAGGA 2592
Db 1 GlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGly 20
QY 2593 TCTGTGGCCATGCAACAGCAAAAGTTCCTCTGAACCAATCGTTTACTGCTCTTACAT 2652
Db 21 SerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuLeuHis 40
QY 2653 GGTTCCTCGATCAGAGATGCTCCATTTTGCACATACCATATATCTAGTATTTTAGTG 2712
Db 41 GlyPheLeuAspGluAsnValHisPheHisThrAsnPheLeuValSerGlnLeuLe 60
QY 2713 AGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGACAGACAGATAGAGTT 2772
Db 61 ArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCys 80
QY 2773 CCTGAATCGGAGAACATTATGACATGTCATCTTTTGGCACTACCTTCAAGAAACCTT 2829
Db 81 ProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 99

RESULT 13
US-11-079-463-7504
; Sequence 7504, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
```

```
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7504
; LENGTH: 624
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-7504

Alignment Scores:
Pred. No.: 5,62e-14 Length: 624
Score: 255.00 Matches: 161
Percent Similarity: 34.1% Conservative: 103
Best Local Similarity: 20.8% Mismatches: 244
Query Match: 4.6% Indels: 267
Gaps: 36

US-10-825-632-2 (1-3120) x US-11-079-463-7504 (1-624)
QY 681 CCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGG 740
Db 1 ProArgTyrTrpAsp-----AsnHisSerLeuLysArg--- 11
QY 741 AGGCCACAAAGGATTTACGCCAACAACTTTAAGGCCCAATCTAGTGGAAATCTAGTTGCC 800
Db 12 -----SerArgThrTyrCysGluLeuThrGluLeuLysThrGlyLysValLeuLeuThr 29
QY 801 CAACATACGATGATCCAAATAATTATGCCGCCCTGATCCAGACTGGATGCTTTTATACA 860
Db 30 AsnLeuArgAspGlyMetArgTrpMetProLysSerAsnLysLeu-----TyrTyrThr 47
QY 861 TAGC-----AACGATATTTGGATATCTAACATCTGTAACCCAGACAGAGAAG 905
Db 48 -ValValAlaProGluGlyAsnAspValIleThrLeuAspProValThrLeuLysGluGl 67
QY 906 GAGACTCACCTATATGTGCACAAATGAGCTAGCCAACTGGAAGAGATGCCAGATCAGCTGG 965
Db 67 uValLeu-----LeuArgGlyIleProGlu----- 75
QY 966 AGTCGCTACCTTTGTTTCTCCAAGAGAATTTGATAGATATCTGCTGATTTGGTGGTCC 1025
Db 76 -----GlnGlyPheSerTrpSerPr 82
QY 1026 AAAAGCTGAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATAT-----GAAGA 1079
Db 82 AsnGluAsp-----PheLeuIleTyrTyrProArgGluGl 94
QY 1080 AATGATGAACTGAGTGGAAATTTATCATGTTATCATCCCT-----ATGTTGGAAC 1133
Db 94 uGlyValLysAspGluGlyProLeuLysArgIleValSerProAlaAspArgIleProAs 114
QY 1134 AAGGAGGCGAGATTCATTC-----CGTTATCTTAAACAGGTACAGCAATCTCTAAAGT 1187
Db 114 nThrArgGlyArgSerPheLeuAlaArgTyrAspIleAlaSerGlyThrSerGluArgLe 134
QY 1188 CACTTTT-----AAGATGTCAGAAATAATGATTTGATGCTGGAAGAGAGAT 1232
Db 134 uThrTyrGlyAsnHisSerThrTyrMetGlnAspIleSerProAspGlyLysTyrLeuLe 154
QY 1233 CATAGATGTCATAGATAAGGAACCTAATCAA---CCTTTTGTAGATT----- 1275
Db 154 uTyrSerSerSerLysGluAsnIleThrGlnArgProPheSerLeuSerSerLeuPheGl 174
QY 1276 -----CTATTTGAAGGAGTTGAATATATTATTCGCCAG 1304
Db 174 nValAsnLeuGluThrLeuAlaValAspThrLeuPhePheGluAspArgPheLeuGlyGl 194
QY 1305 AGCTGGATGGACCTCTGAGGAGGAAAATATGCTTGGTCCATCTCTACTAGATCGCTCCACAG 1364
Db 194 yAlaSerTyr-SerProAspGlyLys----- 202
```



```
Db      83  AspArgGluGlyAspAlaalaGlnLeuTyrIleMetSerThrGluGlyGly-----Glu 100
Qy      1132 ACAAGAGGGCAGATTTCATTCGTTATCTCTAAACACAGGTACAGCAAAATCTCAAGTCACT 1191
Db      101  AlaArgLysLeuThrAspIleProTyr-----GlyValSerLysProLeuTrpSer 117
Qy      1192 TTTAAGATGTCAGAAATAATGATTGAT-----GCTGAAGCAAGGATCATAGATGTC 1242
Db      118  ProAspGlyGluSerIleLeuValThrIleSerLeuGlyGluGlyGluSerIleAspAsp 137
Qy      1243 ATAGATAAG-----GAACTAATTCAACCTTTTGAGATTTCTATTGTAAGAGAGTGA 1293
Db      138  ArgGlnLysThrGluGlnAspSerTyrGluProValGluVal-----GlnGlyLeuSer 155
Qy      1294 TATATTGCCAGAGCTGGATGGACTCCTGAGGGAAAAATATGCTTGGTCCATCCTACTAGAT 1353
Db      156  TyrLysArgAspGlyLysGlyLeuThrArgGlyAlaIleTyrAla----- 169
Qy      1354 CGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGTAATTTATCCCAAGTAGAA 1413
Db      170 -----GlnLeuValLeuValSerValLys----- 177
Qy      1414 GATGATGTTATGGAAGGAGAGACTCATTTAGTACGTGCGCTGATTTCTGTGACGCCACTA 1473
Db      178 -----SerGlyGluMetLysGluLeuThrSerHisLysAlaAspHisGlyAspProAla 195
Qy      1474 -----ATTATCTATGAGAAACA-----ACAGACATCTGG 1503
Db      196  PheSerProAspGlyLysTrpLeuValPheSerAlaAsnLeuLeuThrGluThrAspAla 215
Qy      1504 ATAAATATCCATGACATCTTTCATGTTTCCCAAGTCACGAAGAGGAAATGAGCTT 1563
Db      216  SerLysProHisAspValTyr-----IleMetSer 225
Qy      1564 ATTTTTCCTCTCAATGCAAA-----ACAGGTCTTCCTGATTTATATACAAATATACATCT 1617
Db      226  LeuGluSerGlyAspLeuLysGlnValThrProHisArgGlySerPheGlySerSerSer 245
Qy      1618 ATTTTAAAGGAACCAATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTC 1677
Db      246  PheSerProAspGlyArgTyrLeuAlaLeuLeuGlyAsn-----GluLysGluTyr 262
Qy      1678 AAGTGTCTTATCAAGAGGAGATAGCAATTACAGTGGTGAATGG----- 1722
Db      263  Lys-----AsnAlaThrLeuSerLysAlaTrpLeuTyrAspIleGlu 276
Qy      1723 -----GAAGTTCTTGGCCGGCAT----- 1740
Db      277  GlnGlyArgLeuThrCysLeuThrGluMetLeuAspValHisLeuAlaAspAlaLeuIle 296
Qy      1741 GGATCTAATATCCAGTTGATGAAGTCAGAGCGCTGGTATATTTTGAAGGCACCAAGAC 1800
Db      297  GlyAspSerLeuIleGlyGlyAlaGluGlnArgProIleTrp-----ThrLysAsp 313
Qy      1801 TCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAATCCTGGAGAGGTGACAAGG 1860
Db      314  Ser-----GlnGlyPheTyrValIleGly----- 321
Qy      1861 CTGACTGACCGTGGC----- 1875
Db      322  ---ThrAspGlnGlySerThrGlyIleTyrTyrIleSerIleGluGlyLeuValTyrPro 340
Qy      1876 -----TACTCACATTTCTGTGATCAGTCAGCAGCTGTGACTTCTTT 1917
Db      341  IleArgLeuGluLysGluTyrIleAsnSerPheSerLeuSerProAspGluGlnHisPhe 360
Qy      1918 ATAAGTAAGTATAGTACCAAGAGAGATCCACACTGTGTCTTCCCTTACAGCTATCAAGT 1977
Db      361  IleAlaSerValThrLysProAspArgProSerGluLeuTyrSerIleProLeuGlyGln 380
Qy      1978 CTTGAGATGACCCAACTTCGAAACAAAGGAATTTTGGCCACCATTTTTGGATTCAGCA 2037
```

```
Db      381  GluGluLysGlnLeuThrGlyAlaAsnAspLysPhe----- 392
Qy      2038 GGTCTCTCTCCTGACTATACT-----CCTCCAGAAATTTTCTCTTTTCAAAGTACTACT 2091
Db      393 -----ValArgGluHisThrIleSerIleProGluGluIleGlnTyrAlaThrGluAsp 410
Qy      2092 GGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTCGGAAGAAATAT 2151
Db      411  GlyValMetValAsnGlyTrpLeuMetArgProAlaGlnMetGluGlyGluThrThrTyr 430
Qy      2152 CCTACTGTGCTGTTCATATATGATGCTCTCAGGTGCGATGCTGGTGAATAATCGGTTTAAA 2211
Db      431  ProLeuIleLeuAsnIleHisGlyGlyProHisMetMetTyr----- 444
Qy      2212 GGAAGTCAAGTATTC---CGCTTGATATACCTAGCTCTCTAGGTATATGTGGTCTAGTG 2268
Db      445  GlyHisThrTyrPheHisGluPheGlnValLeuAlaAlaLysGlyTyrAlaValValTyr 464
Qy      2269 ATAGACAACAGGGGATCCTGTCACCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATAAA 2328
Db      465  IleAsnProArgGlySerHisGlyTyrGlyGlnGluPheValAsnAlaValArgGlyAsp 484
Qy      2329 ATGGGTCAAAATAGAAATTGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAT 2388
Db      485  TyrGlyGlyLysAspTyrAspAspValMetGlnAlaValAspGluAlaIleLysArgAsp 504
Qy      2389 GATTTTCATTTGACTTTAGATCGTGGGCATCCAGCGTGGTCTCTATGGAGGATACCTCTCC 2448
Db      505  ProHisIleAspProLysArgLeuGlyValThrGlyGlySerTyrGlyGlyPheMetThr 524
Qy      2449 CTGATGGCATTAATGCAAGGTCAGATATCTTCAGGGTTGCTATTTGCTGGGCCCCCAGTC 2508
Db      525  AsnTrp---IleValGlyGlnThrAsnArgPheLysAlaAlaValThrGlnArgSerIle 543
Qy      2509 ACTCTGTGGATCTTCTAT-----GATCAGGATAC-----ACGGAACGTTAT 2550
Db      544  SerAsnTrpIleSerPheHisGlyValSerAspIleGlyTyrPhePheThrAspTrpGln 563
Qy      2551 ATGGGTCAACCT-----GACCAGAATGAACAGGCTAT 2583
Db      564  LeuGluHisAspMetPheGluAspThrGluLysLeuTrpAspArgSerProLeuLysTyr 583
Qy      2584 TACTTAGGATCTGTGGCGATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTG 2643
Db      584 -----AlaAlaAsnValGlu-----ThrProLeuLeu 592
Qy      2644 CTCTTACATGTTTCTCTGGATGAGAATGTCCATTTTGCATACCATGATATATTACTGAT 2703
Db      593  IleLeuHisGlyGluArgAspAspArgCysProIleGluGlnAlaGlnLeuPheIle 612
Qy      2704 TTTTGTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGC 2763
Db      613  AlaLeuLysLysMetGlyLysGluThrLysLeuValArgPheProAsnAlaSerHisAsn 632
Qy      2764 ATA 2766
Db      633  Leu 633
```

RESULT 15

US-11-079-463-9281

; Sequence 9281, Application US/11079463

; Publication No. US20060073161A1

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR

; FILE REFERENCE: PATH00-03DIV2

; CURRENT APPLICATION NUMBER: US/11/079,463

; CURRENT FILING DATE: 2005-03-14

; PRIOR APPLICATION NUMBER: US 60/128,705

; PRIOR FILING DATE: 1999-04-09

; PRIOR APPLICATION NUMBER: US 09/540,209

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 10444									
; SEQ ID NO 9281									
; LENGTH: 737									
; TYPE: PRT									
; ORGANISM: B.fragilis									
US-11-079-463-9281									
Alignment Scores:									
Pred. No.:	3.56e-09	Length:	737						
Score:	200.50	Matches:	143						
Percent Similarity:	32.9%	Conservative:	107						
Best Local Similarity:	18.8%	Mismatches:	276						
Query Match:	3.6%	Indels:	233						
DB:		Gaps:	31						
US-10-825-632-2 (1-3120) x US-11-079-463-9281 (1-737)									
QY	862	AGCAACGATATTTGGATATCTAATCATCGTAAACAGAGAAGAGAGACTCAGT---	915						
DB	117	AsnArgGluValPheValMetAsnAlaAAspGlyThrAspAsnArgGlnIleThrHisThr	136						
QY	916	---TATGTGCAATAGAG---	930						
DB	137	ProTyrGlnGluAsnGluValThrTrpAlaAlaAspGlySerLysLeuLeuPheLeuSer	156						
QY	931	-----CTAGCCAACATCGAAGAAGATGCCAGATCAGCTGGAGTC	969						
DB	157	AsnAspAsnGlySerSerGlnLeuTyrGluMetAsnProAspGlySerGly---	173						
QY	970	GCTACCTTTTGTCTCCAAAGAAGAATTTGTATAGATAT---	1017						
DB	174	-----ArgLysGlnIleSerLysTyrAspGlyAspIleGluGlyTyr---	187						
QY	1018	TGGTGTCCAAAAGCTGAACAACTCCCAAGTGGTGGTAAATCTTTAGAAATCTATATGAA	1077						
DB	188	-----SerIleSerProAspGlyLysLysIleLeuPheIle---	199						
QY	1078	GAAATGATGATCTGAGGTGGAATATTATCATGTTACATCCCTAGTGTGGAACAAGG	1137						
DB	200	-----AlaGlnValLysThrVal	205						
QY	1138	AGGGCAGATTCTTCGTTATCTCTAAACAGGTACAGCA---	1176						
DB	206	LysSerThrAlaAspLysTyrProAspLeuAspLysAlaThrGlyIleIleIleThrAsp	225						
QY	1177	-----AATCCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGAT	1218						
DB	226	LeuMetTyrLysHisTyrAspGluTrpValThrAlaProHisProPheIleAlaAsp	245						
QY	1219	GCTGAAGGAGG-----ATCATAGATGTCATAGATAAGGAAGTAATTCACACT	1266						
DB	246	PheAspGlyLysSerIleSerAsnIleIleAspValLeuGluGlyGlu-----Pro	262						
QY	1267	TTTGAGATTCTATTGAA-----GGAGTTGAATATATTCGAGAGCTGGATGG---	1314						
DB	263	TyrGluSerProMetLysProTyrGlyGlyIleGluGlnLeuAla-----TrpAsn	279						
QY	1315	ACTCTGAGGGGAAAATATGCTTGGTCCATCTCTACTAGATCGGTCGCCAGACTCGCTACAG	1374						
DB	280	ThrThrSerAspLysValAlaTyrThrCys-----ArgLysLysThrGlyLeuAla	296						
QY	1375	ATAGTGTGATCTCACCTGAATATTATTATCCCAAGTAGAAGATGATGTTATGGAAGGCCAG	1434						
DB	297	TyrAlaIle---	299						
QY	1435	AGACTCATTTGAGTCAGTCGCTGATTCTGTGACGCCCACTAATTAATCTATGAAGAAACA	1494						
DB	300	-----SerThrAsnSerAspIleTyrValTyrAspLeuAsnThr	312						
QY	1495	GACATCTGGATAAATCCATGAC-----ATCTTTGATGTTTTCCTCCCAA-----	1539						
DB	313	LysLysThrValAsnIleThrGluGlyMetMetGlyTyrAspThrAsnProGlnTyrSer	332						

QY	1540	-----AGTCACGAAGAGAAATT	1557						
DB	333	ProAspGlyLysSerIleAlaTrpGlnSerMetGluArgAspGlyTyrGluAlaAspGln	352						
QY	1558	GAGTTTATTTTTCCTCTGAAATGCAAAACAGAGTTTCGTCATTATACAAA-----ATT	1611						
DB	353	AsnArgLeuPheValMetAsnLeuGluThrGlyGluLysArgPheValSerLysAlaPhe	372						
QY	1612	ACATCTATTTTAAAGAAAGCAAAATATAACGATCCAGTGGTGGCTGCCTCCAAGT	1671						
DB	373	AspSerAsnValAspAlaPheValTyrSerArgAspAla-----GGTCAATGG	1722						
QY	1672	GATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATATACCAGT-----	398						
DB	386	-----LysThrIleTyrPheThrGlyValThrPheIleGlyLeuThr	398						
QY	1723	GAAAGTTCTTGGCCCGCATGATCTAATATCAAGTTGATGAAGTCAGAGCGCTGTATAT	1782						
DB	399	GlnIleTyrSerLeuAspLeuSerAsnAspSerValArgProValThrSerGlyMetTyr	418						
QY	1783	---TTTGAAGGCACC-----	1806						
DB	419	AspTyrGluGlyValAlaLeuPheGlyAspLysLeuIleAlaLysArgHisSerMetSer	438						
QY	1807	TTAGAGCATCACCTGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAAGGCTGACT	1866						
DB	439	MetGlyAspGluIleTyrAlaIleAlaLeu-----AspGlyGlnThrThrGlnLeuThr	456						
QY	1867	GACCGTGGCTACTCACATTCTTGTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAG	1926						
DB	456	-----	456						
QY	1927	TATAGTAACGAGAAGATCCACACTGTGTCTCCCTTTACAAGCTACAAGTCCTGAAGAT	1986						
DB	457	-----GlnGluAsnLysGlnIleTyrAspGlnIleGluMetGly	469						
QY	1987	GACCAACTGTGCAAAACAAGGAATTTTGGCCACCATTTTGGATTGACGAGGTCCTCTT	2046						
DB	470	-----LysValGluGlyArgTyrMetLysThrThrAspSerLysGluMetLeu	485						
QY	2047	CCTGACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTGAT	2106						
DB	486	Thr-----TrpValIleTyr	490						
QY	2107	GGGATGCTCTCAAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTC	2166						
DB	491	-----ProGlnPheAspProAsnLysLysTyrProThrLeuLeuPhe	505						
QY	2167	ATATATGTGTGTCTCAGTGCAGTTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTC	2226						
DB	506	CysGluGlyGlyProGlnSerProVal-----SerGlnPheTyrSerTyrArgTyr---	522						
QY	2227	CGCTTGAATATACCTAGCCTCTAGTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2286						
DB	523	AsnMetGlnIleMetAlaAlaAsnGlyTyrIleValValAlaProAsnArgArgGlyLeu	542						
QY	2287	TGTCACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAATAGAAATT	2346						
DB	543	ProGlyPheGlyLeuGluTrpAsnGluAlaIleSerGlyAspTyrGlyGlnCysMet	562						
QY	2347	GACATCAGGTGGAAGGCTCCAAATATCTAGCTTCTCGATATGATTTCATTGACTTAGAT	2406						
DB	563	LysAspTyrPheThrAlaIleAspGluMetAla---LysGluProPheValAspSerAsp	581						
QY	2407	CGTGTGGGATCCACGGTGGTCTTATGGAGGATAC-----	2442						
DB	582	ArgLeuGlyCysValGlyAlaSerPheGlyGlyPheSerValTyrTrpLeuAlaGlyHis	601						
QY	2443	-----CTCTCCCTGATGTCATTAATGACAGGTCAGATATCTCAGGGTTGCTATTGCT	2496						
DB	602	HisAspLysArgPheLysAlaPheIleAlaHisAspGlyIlePheAsnMetGluMetGln	621						
QY	2497	GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACCGAAGCTTATATGGT	2556						

```
Db      622 TyrLeuGluThrGluGluLysTrpPheAlaAsnTrpAspMetGlyGlyAlaTyrTrpGlu 641
Qy      2557 CACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAG 2616
Db      642 LysGlnAsnProThrAlaGlnArgThrPheAlaAsnSerProHisLeuPheValGluLys 661
Qy      2617 TTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCCTGGATGAGAATGTCCAT 2676
Db      662 TrpAspThrPro-----IleLeuCysIleHisGlyGluLysAspTyrArgIleLeu 678
Qy      2677 TTTCACATACCAGTATATTACTGAGTTTTITAGTGGGGCTGGAAAGCCATATGATTTA 2736
Db      679 AlaAsnGlnAlaMetAlaAlaPheAspAlaAlaValMetArgGlyValProAlaGluLeu 698
Qy      2737 CAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTGTAATCGGGAGAACATTATGAA 2796
Db      699 LeuIleTyrProAspGluAsnHisTrpValLeuLysProGlnAsnGlyValLeuTrpGln 718
Qy      2797 CTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATTGTGCTCTTAAA 2853
Db      719 ArgThrPhePheGluTrpLeuAspGlnTrpLeuLysProAsnGluThrAlaGlnLys 737
```

Search completed: April 14, 2006, 12:43:20
Job time : 109.5 secs

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)